ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES

ÉCOLE NORMALE SUPÉRIEURE

Master 2 « Quantifier en Sciences Sociales »
2019-2020

MÉMOIRE DE RECHERCHE

Relations d'amitiés au sein de réseaux sociaux d'une communauté étudiante : Une étude structurale longitudinale

Soutenu par

Mickaël GLAUDEL

Session

Septembre 2020

Sous la direction de

Madame Paola Tubaro
Chargée de recherche, CNRS - Université Paris-Saclay

Rapporteure

Madame Florence Maillochon
Directrice de recherche, CNRS - ENS Paris

Note : ce travail de recherche a été réalisé en conservant une activité professionnelle à plein temps.

Remerciements

Je souhaite adresser de très chaleureux remerciements à madame Paola Tubaro (chargée de recherche-HDR, Laboratoire de Recherche en Informatique, CNRS, Université Paris-Saclay), pour la grande qualité de ses conseils et de ses cours qui m'ont conduit à m'investir avec enthousiasme sous sa direction dans une recherche en sociologie des réseaux sociaux.

Je remercie cordialement madame Florence Maillochon, directrice de recherche, CNRS, ENS Paris, pour avoir accepté d'être rapporteure de mon mémoire de recherche.

Mes amitiés à toute l'équipe du Master « Quantifier en Sciences Sociales » de l'ENS-EHESS, mesdames Florence Maillochon (directrice de recherche, CNRS, ENS Paris), Marie Plessz (chargée de recherche, INRAE, ENS Paris) et Nadine Razgallah (secrétaire, ENS) pour leur cordial accueil.

J'adresse à mon épouse Anne et à ma fille Laure mes affectueuses pensées pour leurs encouragements et leur grande disponibilité durant tout ce temps où mon implication dans notre vie familiale a été mise entre « parenthèse ».

Enfin, toute mon amitié à Marie-Christine.

Sommaire

1. Introduction **4**

2. Revue bibliographique **10**

3. Matériau et méthode **43**

3.1. Formulation des hypothèses et proposition d'opérationnalisation 43

3.2. Présentation de l'expérimentation « Friends and Family » 46

3.3. Positionnement et originalité de la présente étude 57

3.4. Méthodologie pour l'étude descriptive et longitudinale des réseaux 61

3.5. Modélisation longitudinale et dynamique des réseaux 69

4. Résultats **85**

4.1. Analyses descriptives des réseaux uniplexes 88

4.2. Modélisation longitudinale des réseaux uniplexes et multiplexes 108

5. Discussion et conclusions **140**

6. Bibliographie **152**

7. Annexes **160**

8. Table des matières **174**

1. Introduction

L'étude des relations sociales des individus et de leur évolution dans le temps est un large champ de recherche en sciences humaines, sociales et médicales. Ces recherches s'inscrivent dans différentes disciplines, de la sociologie à l'épidémiologie, en ce qui concerne la diffusion des informations et des maladies contagieuses ou encore pour l'étude de l'influence des pairs dans l'adoption d'opinions ou de produits de consommation.

Parmi les multiples phénomènes en jeu dans l'établissement, la rupture et le maintien des relations sociales affinitaires, un des phénomènes le plus robustement établi en sociologie est celui de l'homophilie. Source d'une abondante littérature depuis les travaux de référence de Lazarsfeld et Merton (1954) ou de Maisonneuve et Lamy (1966, 1984) sur les relations amicales, ce phénomène assez intuitif est populairement exprimé par la maxime « qui se ressemble s'assemble ». Celui-ci a des conséquences régulièrement constatées sur les choix des relations amicales et des interactions qui se nouent entre les individus, selon la similarité de leurs attributs socio-démographiques, de leurs opinions, croyances et valeurs. Pour ce qui relève des attributs individuels, il est largement souligné par la littérature scientifique, que les facteurs de genre, d'ethnie, d'âge et de statut professionnel, constituent les principaux vecteurs de l'homophilie dans la sélection par les individus de liens affinitaires tels que l'amitié (McPherson *et al.*, 2001).

Cependant, d'autres facteurs participent à expliquer l'homophilie dans les relations sociales. Ils relèvent des opportunités qui peuvent surgir dans un même espace, physique ou culturel, partagé par les individus (Centola, 2007). Cette homophilie peut alors être considérée comme plus induite par ces espaces que choisie par les individus (McPherson et Smith-Lovin, 1987) : elle est produite par les groupes constitués dans des espaces de socialisation variés qui eux-mêmes ne sont ni neutres ni inertes en ce qui concerne le choix des liens sociaux. En effet, le milieu scolaire, l'environnement professionnel et de loisirs, les structures d'engagements associatifs par exemple, forment des foyers d'activité (« *foci of activity* », Blau, 1977) qui offrent des co-appartenances et des opportunités de rencontres temporaires ou pérennes, plus ou moins opportunes pour l'établissement de liens affinitaires (Fisher, 1977; Mardsen, 1990). Ces espaces constituent des « *structures d'opportunités* » susceptibles de catalyser et d'amplifier le phénomène d'homophilie (McPherson et Smith-Lovin, 1987 ; Stauder, 2014).

Ainsi, le phénomène d'homophilie dans les relations affinitaires s'exprime tout à la fois, d'une part, par une complexe composition d'effets dus aux choix préférentiels des individus et d'autre part, par des effets induits par les groupes et les structures au sein desquelles ces liens naissent, se maintiennent, s'estompent ou meurent.

Au-delà de l'observation du phénomène d'homophilie par des similarités d'attributs individuels et des effets liés aux groupes et aux structures, la sélection relationnelle est également expliquée par la position structurale des individus dans leur réseau social. Une illustration de cette alliance de facteurs socio-démographiques et structuraux nous est donnée, il y a plus de 40 ans, par Eder et Haliman (1978, cités par McPherson *et al.*, 2001, p. 423). A l'issue d'une observation au sein d'une école primaire, ces auteurs constataient que les jeunes filles et garçons développaient des stratégies nettement différentes pour gérer des relations d'amitiés intransitives, c'est-à-dire qui ne relient pas directement le premier et le dernier individu du groupe. Cette situation caractérise un « trou structural », potentiellement instable, qui peut offrir l'opportunité de l'intermédiation par un tiers (Granovetter, 1973 ; Burt, 1982). Ainsi, dans la situation où 3 jeunes d'un trio d'ami·e·s n'étaient pas tous « connectés » entre eux par un lien d'amitié, les filles apparaissaient plus promptes que les garçons à résoudre cette situation intransitive par la rupture d'un des liens d'amitié afin de créer un équilibre. Les jeunes filles résolvaient ainsi « une intransitivité relationnelle » en privilégiant les cliques fermées, alors que les garçons apparaissaient plus « disposés à faire avec » des cercles ouverts d'amis.

Plus généralement, s'il est observé que les individus choisissent de façon préférentielle des pairs qui leur sont similaires sur des attributs individuels socio-démographiques, leurs positions structurales, endogènes au réseau social, jouent également un rôle important dans l'appariement relationnel au sein du réseau (Friedkin, 1993). La littérature des réseaux sociaux montre ainsi que des individus qui sont « structurellement similaires » ont plus de chances d'être en interaction entre eux et que différentes configurations structurales de connectivité relationnelle, telles que la transitivité des relations ou l'intermédiarité, sont des facteurs qui contribuent à expliquer l'évolution du comportement des individus dans l'établissement de liens au sein d'un réseau social (Lorrain et White, 1971; Granovetter, 1973; Holland et Leinhardt, 1976; Burt 1982 ; Lazega, 1988).

Les ressources sociales : un moteur de la dynamique relationnelle

L'homophilie, nous l'avons souligné, est l'un des principaux phénomènes qui facilite l'établissement de relations entre des individus qui partagent des attributs socio-démographiques et des modes vie similaires. Ceci leur permet potentiellement de mobiliser les ressources que peuvent leur procurer leurs semblables et qui nécessitent une confiance réciproque, telles que des services, des conseils, un soutien affectif ou un prêt d'argent. Toutefois, les ressources sociales détenues par des individus semblables tendent à être homogènes et redondantes d'une relation homophile à une autre. L'homophilie peut faciliter et accélérer l'accès aux ressources nécessaires, d'ordre affectif par exemple (Lin, 1982, 1995) lors d'une première phase de constitution d'un groupe et d'intégration des individus nouveaux venus au sein d'une structure, par exemple pour de nouveaux entrants (*freshmen*) à l'Université. Toutefois, l'homogénéité produite peut tout aussi bien être, plus tard, un frein pour accéder à une plus grande hétérogénéité de ressources sociales. Or ces dernières, moins redondantes que celles disponibles au sein d'un groupe homogène peuvent être plus recherchées suite à la phase d'intégration.

En effet, les liens sociaux hétérophiles permettent d'accéder à des ressources différentes, plus rares et potentiellement de plus grande valeur car moins partagées que celles disponibles au sein d'un collectif homogène. Ces ressources, à caractère instrumental (Lin 1982, 1995) contribuent à la satisfaction de besoins plus fonctionnels qu'affectifs, par exemple pour l'atteinte d'objectifs personnels ou professionnels dans une visée à plus long terme que celle de la phase de constitution du groupe. Pour cela, dans le cadre de réflexion de l'acteur rationnel, une stratégie possible est de faire des choix de liens avec des individus dissimilaires, sur des dimensions socio-démographiques et structurales, mais qui potentiellement disposent de ressources absentes au sein du cercle de pairs homogènes (Coleman, 1998, 1990 ; Lin, 1982, 1995 ; Forsé, 1997 ; Flap, 2002). Pour mobiliser ces ressources, il est alors nécessaire d'accéder à ces individus qui diffèrent socialement et qui occupent des positions différentes dans le réseau social considéré. Cette dynamique peut ainsi conduire les individus à un dépassement du choix relationnel homophile vers plus d'hétérogénéité socio-démographique, via une connectivité qui permet d'aller au-delà d'une clique ou d'un groupe d'amis similaires.

A ce jour et à notre connaissance, assez peu de travaux quantifient l'impact sur l'évolution de l'homophilie dans le temps, d'une part, des facteurs socio-démographiques exogènes au réseau qui caractérisent les individus et, d'autre part, des caractéristiques structurales endogènes au réseau.

Ainsi, de nouvelles questions émergent : dans quelle mesure les facteurs structuraux peuvent-ils moduler, amplifier ou atténuer les facteurs socio-démographiques observés dans le phénomène d'homophilie ? Dans l'expression de l'homophilie observée sur l'appariement affinitaire, quel est le poids de la position structurale d'un individu comparé à ses caractéristiques socio-démographiques ?

L'entrelacement des relations sociales : un récit de la sociabilité

L'étude des formes diverses de contacts entre les individus montre que les caractéristiques de l'homophilie tendent à se renforcer avec la multiplication des relations entre les individus, selon un effet cumulatif (Fisher, 1982 ; Welleman *et al.*, 1966). Ainsi, les liens sociaux multiples tendent à révéler une homophilie plus prégnante que s'ils sont appréhendés sur une seule dimension. Toutefois, McPherson *et al.* (2001) font le constat qu'assez peu d'études éclairent les impacts sur l'expression de l'homophilie dans les relations, d'une part, de la multiplexité des relations et, d'autre part, de l'évolution dans le temps de ces réseaux multiplexes de relations. Or, comme le note Alain Degenne (2013, p.38) : *« les relations et donc les réseaux sont des histoires, il faut donc les analyser comme tels : la prise en compte du multi-niveaux, cercles et relations, est importante »*.

En effet, les différentes modalités d'interactions sociales entre les individus, les échanges en face-à-face, par téléphone ou courriel par exemple, « *s'entrelacent étroitement* » et se cumulent selon l'intensité de la relation affinitaire (Smoreda *et al.*, 2007 ; Cardon et Smoreder, 2014). Ainsi, « *plus on se sent proche de quelqu'un, plus on convoque une grande variété de dispositifs d'interactions* » (Bailliencourt *et al.*, 2007, p.112). Par conséquent, afin d'appréhender et modéliser le comportement des individus et l'évolution de leurs relations sociales au sein d'une communauté, il apparaît important de prendre en compte le plus possible l'ensemble des actes de communication qu'ils y déploient simultanément et qui narrent une histoire qui s'écrit par l'entrelacement de modalités multiplexes d'échanges.

Des questions supplémentaires émergent alors : qu'en est-il de l'impact sur le phénomène d'homophilie de la diversité des vecteurs d'échanges entre les individus ? Une rencontre régulière en face-à-face entre deux individus témoigne-t-elle du même phénomène d'homophilie qu'un échange régulier par téléphone ?

Une question se pose également sur l'écart entre les liens affinitaires « déclarés » et les traces observables des échanges sociaux entre les individus. Ainsi, les traces d'échanges, en intensité et dans le temps, entre les individus convergent-elles avec la perception par les individus de leurs liens affinitaires ? Dans quelle mesure les comportements des individus, observés par l'intensité de leurs rencontres et de leurs contacts téléphoniques, traduisent-ils les préférences affinitaires déclarées par les individus ? Quelle est l'évolution conjointe dans le temps des préférences affinitaires déclarées par les individus et de l'observation des traces de leurs échanges en présentiel et en distanciel ?

Une exploration longitudinale de l'homophilie au sein de réseaux sociaux multiplexes

Les réseaux sociaux formés durant les temps longs d'études des individus présentent l'avantage de permettre d'observer de façon instantanée ou longitudinale et de disséquer les phénomènes en jeu dans l'établissement des relations affinitaires d'un public captif considéré comme un réseau complet, au sein d'un environnement certes spécifique mais circonscrit. Ceci peut être réalisé par des observations directes, des enquêtes déclaratives uniques ou répétées dans le temps (Moreno, 1934 ; Russell-Bernard *et al.*, 1979 ; Verbrugge, 1977), par la mise en œuvre d'outils de collecte de données *via* des capteurs de co-présence, des téléphones portables ou encore par l'exploitation de traces d'interactions sur l'Internet (Kossinets et Watts, 2009 ; Steglich *et al.*, 2010 ; Aharony *et al.*, 2011 ; Smieszek *et al.*, 2011 ; Stehlé *et al.*, 2013 ; Fournet et Barrat, 2014). Parfois, plusieurs de ces modalités de recueil d'observations des interactions sociales sont utilisées simultanément afin de restituer les dimensions de la sociabilité des individus.

Une opportunité d'apporter des éléments de réponse aux questions soulignées ci-dessus, nous est fournie par l'exploitation secondaire de données collectées dans le cadre d'une expérimentation menée par l'Institut Technologique du Massachussets (MIT). Au cours de l'année universitaire 2010-2011, le MIT a conduit une ambitieuse expérimentation basée sur le suivi longitudinal d'une centaine de jeunes professionnels de provenance géographique variée en reprise d'études (Aharony *et al.*, 2011). Cette expérimentation, nommée « *Friends and Family* », consistait à suivre durant une année les interactions sociales quotidiennes d'une centaine d'étudiants au sein d'une communauté résidentielle étudiante familiale du MIT. Un des intérêts de cette communauté est d'être composée d'étudiants de huit origines ethniques.

A l'aide d'un important dispositif technologique, parmi les nombreuses données collectées, les co-présences, les contacts téléphoniques et la perception des liens d'amitiés entre ces individus ont été recueillis régulièrement au sein de plusieurs bases de données relationnelles.

Par une analyse structurale longitudinale des agencements uniplexes et multiplexes des réseaux sociaux que nous nous proposons de reconstruire à partir des données mises à disposition par le MIT, nous souhaitons répondre aux questions énoncées précédemment et synthétisées comme suit :

1. Dans quelle mesure le phénomène d'homophilie contribue-t-il à expliquer l'évolution de la structure des relations amicales des individus au sein d'une communauté ? Comment les caractéristiques socio-démographiques et structurales propres à l'homophilie co-évoluent-elles dans le temps ?
2. De quelle façon l'expression de l'homophilie se révèle-t-elle, d'une part, dans les réseaux uniplexes et, d'autre part, dans les réseaux multiplexes ?
3. Dans quelle mesure les comportements des individus, observés par l'intensité de leurs rencontres et de leurs contacts téléphoniques, traduisent-ils des préférences affinitaires déclarées par les individus ?

Nous prenons en compte le caractère multi-ethnique de la communauté suivie, en portant un regard particulier sur les différentes formes d'expression de l'homophilie relative à l'origine ethnique des individus.

Nous proposons d'exposer au chapitre 2 notre démarche d'investigation en débutant par une revue bibliographique qui sert de support à la construction de notre objet d'étude. Nous présentons ensuite au chapitre 3 la déclinaison de notre problématique par la formalisation d'un ensemble d'hypothèses et une opérationnalisation. Nous détaillons également au chapitre 3 le matériau exploité et la méthode adoptée. Enfin, nous présentons nos résultats au chapitre 4 et nous proposons une discussion conclusive au chapitre 5.

2. Revue bibliographique

« Les relations émergent, évoluent, s'étiolent, se rompent, se renouent, se transforment, changent de tonalité » (Bidart et al., 2011, p.12).

Dans ce chapitre, à travers une revue bibliographique de travaux en sociologie générale et sociologie des réseaux, nous présentons tout d'abord le phénomène de l'homophilie ainsi que ses principaux facteurs et impacts dans la sélection des relations amicales et affinitaires entre les individus. Les termes de relations amicales et affinitaires sont précisés. Ensuite, nous examinons en quoi les concepts d'acteur rationnel et de capital social peuvent contribuer à éclairer l'évolution de la création, du maintien et de la rupture des relations au regard du phénomène de l'homophilie. Enfin, nous soulignons l'intérêt de prendre en compte les aspects temporels et la multiplexité des modalités d'interactions entre les individus afin d'appréhender la complexité d'un possible récit de leur sociabilité.

Sociabilité et réseau social

Tout d'abord, le terme « sociabilité » nécessite quelques précisions quant au point de vue considéré dans cette étude ainsi que certains des concepts utilisés issus de l'analyse structurale des réseaux sociaux.

L'utilisation du terme de sociabilité peut faire référence à des liens sociaux de nature extrêmement variée, tels que les liens de parenté, d'amitié, d'amour, de travail ou de voisinage (Héran, 1988). Ce terme fait également communément référence à un jugement sur l'aptitude d'un individu à nouer des relations avec autrui, d'un point de vue d'ordre psychologique. Notre étude s'inscrit dans le champ de la sociologie, nous n'investissons donc pas la question des aptitudes cognitives et affectives d'un individu à lier une relation amicale avec des pairs. Nous adoptons dès lors une définition simple de la sociabilité comme « l'ensemble des relations qu'ils [les individus] entretiennent avec d'autres individus » (Degenne et Forsé, 2004, p.35). Nous considérons que cette sociabilité s'incarne dans des échanges relationnels qui évoluent au cours du temps et au sein d'une structure sociale donnée.

Par conséquent, dans le cadre de notre étude nous appréhendons la sociabilité des individus à partir, d'une part, du constat que des individus déclarent des relations affinitaires et interagissent entre eux selon des modalités d'échanges variées et, d'autre part, de l'hypothèse centrale que la dynamique de cette sociabilité - restituée et observée - peut être modélisée dans un contexte et un temps circonscrits.

La sociabilité étant ainsi précisée, la question est ici d'interroger comment et pourquoi se forment, se déforment et s'estompent les relations amicales entre des individus au sein d'un réseau social et au cours du temps. Afin d'apporter des éléments de réponse à cette question centrale, nous mobilisons les concepts et les outils de l'analyse structurale proposés par la sociologie des réseaux sociaux. Cette approche conceptuelle et méthodologique contribue à mieux comprendre les propriétés structurales d'ensembles sociaux, c'est-à-dire l'évolution des « interdépendances et absences d'interdépendances entre les membres d'un milieu social organisé » (Lazega, 2014, p. 3). A partir de données relationnelles, l'analyse des réseaux sociaux permet de rendre compte de l'importance des caractéristiques des individus mais également de la structure formée par leur relation, dans l'étude de l'évolution de leur liens sociaux.

Le champ de l'analyse des réseaux sociaux est riche de métaphores telles que liens, nœuds, trous, ponts, fermeture. Elles apportent une puissance évocatrice très utile pour appréhender « l'insaisissable lien » entre le choix des individus et les structures des réseaux de sociabilité (Favereau et Lazega, 2002, p.11). Cela participe à la lisibilité sociologique des phénomènes mis en lumière par les concepts et les outillages issus de la théorie des graphes (Berge, 1958) que nous mobilisons. Toutefois, comme le soulignent Degenne et Forsé (2004), afin de dépasser le champ de la métaphore il convient de se doter de mesures opérationnelles.

L'étude de la sociabilité au sein d'un réseau social

Le réseau social évoqué dans notre questionnement peut être désigné diversement selon les caractéristiques de la structure sociale observée et l'angle d'observation adopté. Il peut s'agir d'observer les relations développées par un individu que l'on distingue au centre d'un réseau de liens sociaux. On parlera dans ce cas de réseau « Ego centré ». Si l'on considère un réseau social qui comprend les liens entre un ensemble d'individus, il s'agira alors d'un « réseau complet ». En ce qui concerne la nature des liens on distinguera : leur caractère formel ou informel, les modalités uniques ou multiples d'échanges, noués en présentiel ou à distance ainsi que la nature de leur restitution (déclarés, remémorés ou observés). La combinaison des angles d'observations, de la nature des liens, de leur caractère uniplexe ou multiplexe et de l'environnement considéré, ouvre la possibilité d'études sur des champs variés : l'évolution des relations au sein d'une communauté d'étudiants, d'un groupe d'amis, d'une association, d'une entreprise ou d'un forum d'échanges sur Internet.

Il pourra s'agir, par exemple, d'étudier les relations d'amitiés entre des jeunes filles dans un pensionnat de rééducation (Moreno, 1934), les liens entre des habitantes d'une petite ville du sud des Etats-Unis (Davis *et al.*, 1941), le quotidien relationnel des habitants d'une île (Barnes, 1954), l'évolution des liens d'amitié ou d'animosité entre des moines (Sampson, 1969) ou entre des étudiants d'une fraternité (Russell *et al.*, 1979-1980), l'évolution des relations d'un cercle d'amis du lycée à la naissance du premier enfant (Bidart, 1997, 2011) ou des échanges de conseils entre des juges d'une cour de justice (Lazega *et al.*, 2012). Les terrains d'études sociologiques sont nombreux et les approches méthodologiques diffèrent selon les contextes sociaux étudiés, le type de réseau social observé (centré sur un individu ou sur un groupe complet) et la temporalité considérée (une observation « en coupe » ou plusieurs observations longitudinales).

Les fondements de l'analyse structurale des relations sociales sont consensuellement attribuées aux travaux de Georg Simmel et de Célestin Bouglé (19ème siècle) qui formeront les éléments du cadre conceptuel de l'interactionnisme structural. Simmel explicite ainsi l'idée que « les groupes dont l'individu fait partie constituent un système de coordonnées tel que chaque coordonnée nouvelle qui vient s'y ajouter le détermine de façon plus nette » (cité par Degenne, 2013, p.24). Dans ce cadre conceptuel, un individu est présenté comme pouvant être caractérisé par l'ensemble des relations sociales qu'il développe avec autrui et qui forme ainsi une structure sociale. Quelques décennies plus tard, les études sociométriques développées par Moreno (1934) montrent les potentialités de l'analyse des structures formées par les relations entre des individus. Ces travaux contribueront à l'émergence de la sociologie des réseaux sociaux.

Depuis les travaux de Simmel, Bouglé, Moreno et de Peter Blau (1964) notamment, l'étude systématique de l'évolution des relations sociales entre des individus a connu un regain d'intérêt dans la seconde moitié du 20ème siècle avec le développement des concepts et méthodes de la théorie des graphes développée notamment par le mathématicien Claude Berge (1958). Depuis quelques décennies, la mise en œuvre des outils analytiques issus de la théorie des graphes a bénéficié de l'évolution de la rapidité de calcul des outils informatiques, de leur capacité à traiter des données de nature extrêmement variée dont le volume s'accroit de façon exponentielle, ainsi que du développement d'algorithmes de plus en plus efficients (Abiteboul, 2012 ; Mayer-Schöberger et Cukier, 2014 ; Cardon, 2015). Les développements théoriques, les progrès techniques et la disponibilité croissante de bases de données relationnelles ont stimulé un regain de popularité de l'analyse structurale en sociologie.

La disponibilité des données relationnelles longitudinales et des outils d'analyses variés nous permettent d'adopter cette approche afin d'investir de façon systématique et temporelle les relations uniplexes et multiplexes nouées et dénouées entre des individus, par des modélisations comprenant tout à la fois des facteurs exogènes et endogènes à la structure formées par ces relations.

L'homophilie et la sélection des relations sociales

« The old proverb says that "birds of a feather flock together", similarity begets friendship ».

La littérature sur le phénomène de l'homophilie dans les liens sociaux est riche et multidisciplinaire. Elle porte sur des contextes sociaux variés : les engagements associatifs et politiques, les liens professionnels, les relations affinitaires d'amours et d'amitiés. Parmi les principaux phénomènes qui contribuent à l'élaboration d'une relation amicale, l'importance de l'homophilie est largement confirmée par des travaux de recherche variés, tels que ceux menés sur l'amitié par Lazarsfeld et Merton (1954) et Maisonneuve et Lamy (1993). Il s'agit d'un invariant résumé par le proverbe « qui se ressemblent s'assemblent » ou « *similarity breeds connection* ». Le phénomène d'homophilie est décrit comme la forte tendance des relations sociales, dont particulièrement celles du registre de l'amitié, à se former entre des individus qui possèdent des caractéristiques socio-démographiques des attitudes, des croyances et des centres d'intérêts similaires plutôt qu'avec des individus dissimilaires. Ce phénomène participe à structurer les réseaux sociaux d'amis, c'est-à-dire la création, le maintien et la dissolution des liens d'interaction entre les individus.

McPherson et Smith-Lovin (1987) et McPherson *et al.* (2001) soulignent les principales modalités de ce phénomène. Ainsi, ils distinguent l'homophilie de base (*baseline homophily*), l'homophilie consanguine (*inbreeding homophily*) en contraste avec l'homophilie non consanguine (*outbreeding homophily*). Selon les auteurs, ces modalités conditionnent le degré d'homophile observée dans les appariements dyadiques. La première modalité, l'homophilie « de base », est la plus attendue. D'une part, celle-ci est liée aux attributs socio-démographiques des individus dont principalement l'âge, l'ethnie, le genre et le niveau de diplôme. D'autre part, elle relève également des attitudes et des comportements des individus. La deuxième modalité, l'homophilie consanguine, est observée au-delà de l'homophilie de base attendue. Elle est liée aux préférences personnelles qu'expriment les individus dans leurs choix d'appariement avec leurs pairs.

Les auteurs apportent un raffinement de ces concepts avec l'introduction de l'effet dû à la composition du groupe (ou structure) au sein duquel les individus sélectionnent leurs relations. Ainsi, la prise en compte de l'influence de la composition du groupe comparée à la composition de la population dans son ensemble, permet de distinguer une homophilie induite *(induce homophily*) et d'une homophilie choisie (*choice homophily)*. L'homophilie induite est produite par le groupe alors que l'homophilie choisie est produite par les choix individuels.

Dans un premier cas, si l'on considère que la composition du groupe est un fidèle reflet de l'hétérogénéité observée au sein de la population globale, le groupe ne constitue pas une structure susceptible d'offrir des opportunités de produire d'autres appariements que ceux qui sont observés dans la population (McPherson et Smith-Lovin, 1987). Le groupe en tant que structure, par exemple une association ou une école, est considéré comme un « creuset local » (*local arena*, p.371) de la population globale : la sélection des liens affinitaires s'opère avec la même similarité que celle observée dans la population entière. Dans un second cas, la composition du groupe est strictement homogène ce qui n'est pas le cas pour la population globale. Le groupe va alors contraindre la similarité des appariements entre les individus, ce qui produira des dyades homophiles. Cette homophilie est alors induite par le groupe au sein duquel les liens sociaux se nouent. Le groupe, en tant que structure, fournit les opportunités de création des liens. McPherson *et al.* (2001) soulignent que l'homophilie consanguine n'est pas équivalente à l'homophilie choisie : elle peut tout à la fois être induite par les structures sociales et par les préférences personnelles. Les auteurs relèvent que des facteurs structuraux tels que la densité des relations et l'encastrement (*embedenness*), interviennent dans l'expression de l'homophilie et complexifient ce qui pourrait apparaitre comme une opposition binaire et disjointe entre des modalités d'homophilie prises deux à deux.

Par ailleurs, dans leurs recherches sur les processus de relations amicales Lazarsfeld et Merton (1954) distinguent l'homophilie de statut de l'homophilie de valeur. Selon les auteurs, l'homophilie de statut correspond à la similarité des statuts formels et informels des individus. Les statuts sont soit du registre des attributs dits « prescrits » tels que l'ethnie, l'âge ou le genre, soit du registre des attributs dits « acquis » tels que la religion ou l'éducation. L'homophilie de valeur est plutôt présentée comme une similarité de croyances qui influencent le comportement de l'individu.

Une autre distinction des modalités d'expression du phénomène d'homophilie dans les relations sociales est apportée par Van Duijn *et al.* (2003). A l'occasion d'une étude sur les réseaux d'amitiés formés par des étudiants nouveaux arrivants (*freshmen*) observés durant plusieurs mois au sein d'une Université des Pays-Bas, les auteurs distinguent deux catégories de similarités dyadiques. Ils qualifient ces similarités de « visibles » et d'« invisibles » (p.165). Les similarités visibles correspondent à des attributs socio-démographiques, tels que l'âge et le genre. Ces similarités désignent un ensemble hétérogène d'activités, telles que le sport, une sortie au cinéma ou un diner au restaurant, propices à l'émergence de relations sociales. Elles désignent aussi des attitudes et comportements ainsi qu'un ensemble d'« indicateurs d'une sous-culture chez les jeunes » (p.166) tels que la consommation de drogue douce ou encore les échanges de confidences (*confiding matters*).

Les échanges de confidences sont également retenus par Marsden (1981, 1988) comme marqueurs de l'homophilie dans les relations sociales. Pour l'auteur, l'ethnie, le type d'activité professionnelle ou encore la religion sont des variables qui stratifient et ferment, d'une part, les relations entre les groupes d'individus et, d'autre part, les relations au sein des groupes. Ce type de variable génère une tendance homophile à créer ou non un lien de confiance entre les individus, mesuré par la nature du contenu des discussions. Ainsi, à partir données d'enquêtes réalisées auprès de 4000 individus de nationalité américaine, l'auteur s'intéresse à l'homogénéité dans les relations sociales mesurée par la fréquence des discussions sur des sujets importants et des confidences intimes (*confiding relations*). En exploitant ce matériau et en adoptant le cadre théorique de Blau (1977), que nous commenterons ci-après, Marsden distingue les caractéristiques « nominales » de l'homophilie de ses caractéristiques « hiérarchiques ». Ces deux caractéristiques œuvrent différemment dans la structuration des liens sociaux. Les variables hiérarchiques continues (l'âge) ou discrètes (le prestige de la profession ou le niveau de diplôme) participent à l'émergence d'une différenciation inter-groupe et intra-groupe selon une distance sociale (*social distance differenciation*, p.57). Les variables nominales (l'ethnie, le genre) quant à elles ont tendance à densifier l'homogénéité intra-groupe. Verbrugge (1977) étend les similarités des individus à un effet relatif à leur « ronde quotidienne » (*daily rounds*, p. 577), c'est-à-dire leurs routines de déplacements liées à l'homogénéité de statut social et spatial. L'auteur souligne que les individus qui partagent les mêmes rôles sociaux et croyances ont plus de chances de se croiser dans leurs déplacements pour le travail et les loisirs et donc d'interagir et créer des liens.

Relation affinitaire et ethnie

Plusieurs études citées précédemment mentionnent l'importance majeure de l'origine ethnique[1]/race dans l'établissement des relations sociales. L'homophilie est observée comme un phénomène particulièrement prégnant lors de la prise en compte de cet attribut dans les relations amicales, scolaires, professionnelles et intimes. Ainsi, Mc Pherson et Smith-Lovin (1987, 2001) soulignent que « La race et l'ethnie sont clairement les plus grands diviseurs dans les réseaux sociaux aujourd'hui aux Etats-Unis » (2001, p.420).

Ibarra (1995) explore cette dimension dans une recherche sur les liens entre l'origine ethnique et les opportunités professionnelles. L'auteur étudie les réseaux sociaux professionnels formels et informels d'une soixantaine de managers de quatre grandes entreprises aux Etats-Unis. Ceux-ci sont d'origine ethnique blanche, afro-américaine, hispanique et asiatique. Au sein de cette population, le facteur relatif à l'origine ethnique explique les différences de progression de carrière entre des managers de niveaux de performance et de compétences similaires. L'auteur souligne plusieurs éléments saillants. Les managers d'origines ethnique minoritaires et ceux d'origine ethnique blanche, tout en exerçant un travail similaire avec un statut comparable, ont dans leurs réseaux professionnels des liens très différents du point de vue de l'homophilie et de l'intimité. Ainsi, il apparaît que les managers d'origines ethniques minoritaires sont conduits à chercher des relations professionnelles amicales de même origine ethnique au-delà de leur milieu professionnel proche et de leur niveau statutaire. Ceci a pour conséquence qu'ils bénéficient de réseaux de conseils et de soutien bien plus hétérogènes qualitativement que ceux des managers d'origine ethnique majoritaire. Enfin, des différences de niveaux hiérarchiques et de multiplexité apparaissent entre les liens des réseaux professionnels des managers à hauts potentiels d'origines ethniques minoritaires et ceux d'origine ethnique blanche. Pour l'auteur, ces éléments soutiennent l'évidence d'une « contrainte structurale » effective dans ces réseaux professionnels (p.693).

[1] De nombreux auteurs cités exploitent des données issues d'enquêtes réalisées aux Etats-Unis d'Amérique et utilisent conjointement, de façon indifférenciée, les termes de « race » et d'« ethnie » en se référant aux catégories suivantes : « blanc, noir, hispanique, afro-américain, asiatique… ». Par exemple : « *interracial/interethnic* » (Vaquera et Kao, 2008) ou encore « *race/ethnicity* » (Mc Pherson et Smith-Lovin, 2001; Aharony *et al.*, 2011) et « *race* » (Moody, 2001). L'argumentation permettant de différencier « la race » de « l'ethnie » dépassant le cadre de cette étude, nous faisons le choix de traduire les termes utilisés par les auteurs. Toutefois, dans la suite de ce mémoire nous utiliserons le terme « origine ethnique ». Le lecteur intéressé pourra lire les actes de la journée d'étude des doctorants de Mondes Américains (novembre 2014), « Race, ethnie, identité : écrire une histoire hémisphérique des signifiants identitaires », coordonnée par Heredia Neyra J.J. et Ollivier S. (https//journals.openedition.org/nuevomundo).

A partir de l'exploitation de 1 640 profils d'étudiants sur la plateforme *Facebook,* Wimmer et Lewis (2010) explorent les liens d'amitiés plus ou moins forts entre les individus d'origine ethnique asiatique, blanche, noire et hispanique. Les auteurs mettent évidence d'une part, une homophilie entre groupes relative à la race/origine ethnique et, d'autre part, une homophilie secondaire intra-groupe, qu'ils qualifient d'« homophilie micro-ethnique » (p.612), identifiable au sein des groupes de même origine ethnique (*coethnics*). Par une analyse structurale, les auteurs soulignent également l'importance de facteurs structuraux, tels que la transitivité, qui amplifient l'effet d'homogénéisation. Par ailleurs, le niveau socio-économique des étudiants et la proximité spatiale des lieux de vie contribuent également fortement à l'homophilie observée.

Sur ce même thème, Moody (2001) explore les liens entre l'hétérogénéité raciale, l'intégration scolaire et la ségrégation dans les relations affinitaires aux Etats-Unis à partir d'un échantillon comprenant une centaine d'écoles représentatives (classe d'âge de 7 à 12 ans). Les résultats de ces travaux montrent une forte relation positive entre l'hétérogénéité raciale et la ségrégation dans les liens affinitaires. Ainsi, l'auteur souligne que le haut niveau d'hétérogénéité raciale qui résulte d'une organisation scolaire qui vise la réduction des ségrégations raciales ne favorise pas en fait l'intégration raciale. Au contraire, l'auteur met en évidence « qu'en structurant les contacts inter-raciaux, l'organisation scolaire influence la ségrégation raciale dans les liens d'amitié » (p.709). L'effet le plus remarquable sur la réduction de la ségrégation provient des activités « hors-programme » (*extracurricular*) organisées par les écoles. Plus ce type d'activités parvient à mélanger les origines raciales des élèves, moins il est observé de ségrégation dans les liens d'amitiés. Il apparaît que lorsque des élèves hétérogènes sur la dimension raciale se rencontrent au sein de foyers d'activités propices à offrir des opportunités de rencontres et d'actions collectives vers un but commun, dans une relation égalitaire, l'intégration fournit des résultats probants.

Toutefois, Moody (2001) souligne que des études longitudinales sur l'évolution des liens d'amitié sont nécessaires pour explorer ce phénomène, par exemple en ce qui concerne la stabilité des liens amicaux entre des individus de race similaire. En effet, la qualité comparée des liens homogènes et hétérogènes selon cet attribut demeure assez méconnue.

Par exemple, Vaquera et Kao (2008) soulignent que « si d'une part, théoriquement on peut s'attendre à ce que les relations inter-raciales/inter-ethniques soient plus ténues que les relations intra-ethniques, selon le principe normatif d'homophilie [observée sur cette dimension aux Etats-Unis], d'autre part, les relations inter-ethniques étant plus difficiles à nouer compte-tenu des pressions sociales, on peut également s'attendre à ce que ces amitiés hétérogènes soient plus intimes une fois nouées » (p.5). A partir de l'exploitation de données issues d'une enquête nationale sur la santé de 90 000 adolescents aux Etats-Unis, pays dans lequel l'homophilie d'ethnie est encore très présente au sein des écoles américaines 70 ans après le « *Brown vs Board of Education act* », les auteures observent que les liens inter-ethniques sont moins réciproques que les liens intra-ethniques et qu'ils sont en outre plus difficiles à maintenir dans le temps et finalement moins intimes. En examinant les relations au sein des groupes ethniques minoritaires, les résultats de l'étude montrent des différences significatives dans la formation et le maintien des liens d'amitiés selon les différentes ethnies des dyades observées.

Plusieurs éléments d'interprétation sont proposés, telles que les différences entre des cultures plus collectivistes où la relation dyadique est moins valorisée que l'appartenance au groupe et des cultures plus individualistes où le lien avec le meilleur ami prime sur le lien avec le groupe. Toutefois, les différences de perception des liens d'amitié chez les jeunes selon les ethnies, le volume de données manquantes sur la réciprocité des liens de niveau « meilleur ami », le caractère non-longitudinal de l'étude et l'absence de prise en compte d'autres effets structuraux que la réciprocité, limitent la portée des conclusions, notamment sur la stabilité et la durabilité des liens observés.

La communauté d'individus dont nous proposons d'explorer l'évolution des relations d'amitié est composée par une centaine d'étudiants avec huit origines ethniques (amérindienne, asiatique, blanche, hispanique, israélienne, noire, moyen-orientale, métissée). Cette communauté est ainsi « multi-ethnique » en écho à l'attractivité mondiale attendue pour le MIT, 4ème université mondiale dans le classement de « Changaï » pour l'année 2020[2]. Par conséquent, dans notre étude nous faisons le choix de privilégier cet attribut parmi les autres attributs individuels disponibles (le sexe et le nombre d'enfants).

[2] Le classement de « Shanghai » réalisé par l'« *Academic Ranking of World Universities* » désigne depuis 2003 les 1 000 meilleures universités au monde. Données publiées le 15/08/2020 (www.shanghairanking.com).

Homophilie et structures sociales

Plus travaux évoqués précédemment mentionnent que l'expression du phénomène d'homophilie ne s'explique pas de façon complètement satisfaisante uniquement par des similarités d'attributs. En effet, les questions soulevées par l'étude de l'homophilie dans les relations affinitaires dépassent la similarité des attributs acquis et hérités, des statuts et des valeurs des individus. L'appréhension du phénomène interroge également le rôle des structures sociales *via* lesquelles les individus se rencontrent et des opportunités qui en émergent. Ces structures offrent des opportunités de contacts entre les individus et contribuent à expliquer comment et pourquoi la rencontre d'individus (*meeting)* évolue vers un appariement (*mating*) entre certains d'entre eux.

Par conséquent, il convient d'explorer les propositions de réponse à la question suivante : comment les structures sociales participent-elles à expliquer le phénomène d'homophilie dans l'appariement affinitaire des individus ? En réponse à cette question, Blau (1977) et Feld (1981) proposent des cadres conceptuels qui contribuent à éclairer le rôle des structures sur l'appariement préférentiel des individus. Le premier auteur développe une théorie macro-structurale des structures sociales (*macrosociological theory of social structure*, Blau, 1977) et le second une théorie de la focalisation des liens (*focus theory*, Feld, 1981).

En posant les bases d'une sociologie macrostructurale, Blau (1977) entreprend d'analyser les relations entre des groupes considérés comme des « sous-unités » d'une population globale. Dans sa théorie Blau présente les structures sociales comme « des distributions de la population parmi différentes positions sociales qui reflètent et impactent les relations entre les individus » (p.28). Les individus par leur appartenance à des groupes religieux ou professionnels, leur niveau de diplôme ou leur origine ethnique, occupent simultanément différentes positions sociales selon une certaine distribution. Les positions des individus sont ici leurs attributs et non leurs positions dans leurs relations sociales. Ces positions sont établies selon des « paramètres nominaux » tels que la religion et l'ethnie et des « paramètres continus » tels que le niveau d'études ou de salaire. Pour Blau (1977), le niveau de différentiation nominal reflète l'hétérogénéité, c'est-à-dire la distribution de la population dans des groupes selon les paramètres nominaux. Les groupes rassemblent des individus qui partagent un attribut nominal donné qui influence les relations qu'ils nouent entre eux, sans considération de rang. Le niveau de différentiation continue, quant à lui, différentie les individus selon un rang qui influence leurs relations et qui reflète l'inégalité.

L'hétérogénéité et l'inégalité sont les deux formes de différentiation dans la structure sociale définie par Blau (1977) : l'hétérogénéité est la probabilité que deux individus de la population appartiennent à des groupes différents, tandis que l'inégalité est formée par la distance de statut entre deux individus. La macrostructure des sociétés est alors définie comme l'espace multidimensionnel des positions sociales parmi lesquelles les individus sont répartis et qui influencent leurs relations.

Dans ce cadre théorique, l'étude des macrostructures porte sur les modèles de relations sociales entre des positions similaires ou différentes occupées par les individus, sans disséquer les réseaux des relations entre eux. Ainsi, pour Blau, « l'analyse des réseaux considère les nœuds selon les liens qui les relient et contrastent leurs rôles, similaires ou différents, selon une équivalence structurale » alors que « l'analyse macrostructurale considère, elle, les variations parmi les positions des individus » (Blau, 1993, p.204). Pour l'auteur, les structures sociales ne prédéterminent pas l'action des individus : d'une part, elles constituent des opportunités pour les individus à nouer des liens et, d'autre part, elles exercent « une contrainte en limitant les opportunités » à nouer des liens (Blau, 1993, p.204). En résumé, les structures sociales fournissent des conditions favorables à la rencontre et à l'appariement et elles agissent de façon contraignante sur les opportunités d'appariement entre les individus.

Par une théorie de focalisation des liens (*focus theory)*, Feld (1981, 1982) propose une approche de l'organisation sociale des liens d'amitiés qui invite au dépassement de la seule prise en compte des attributs individuels dans l'exploration des mécanismes de l'homophilie.

A partir de l'exploitation de jeux de données sur les relations amicales de plus de 300 travailleurs d'une usine de Détroit et d'un échantillon de plus 1 000 résidents du nord de la Californie, Feld (1982) observe que l'homogénéité des foyers étudiés est un facteur qui contribue de façon conséquente à expliquer l'homophilie d'âges, d'attitudes, de comportements et de statut social. Les foyers d'activités étudiés, en tant que structures sociales, ne sont cependant pas suffisants pour expliquer l'homophilie. Toutefois, les variations d'homogénéité au sein de ces foyers apparaissent avoir un large effet sur les variations constatées dans le phénomène de l'homophilie. Pour l'auteur, « négliger [la prise en compte] des facteurs structuraux a conduit à surestimer l'importance de l'appariement préférentiel des individus avec ceux qui leur ressemblent » (Feld, 1982, p.797). Ainsi, si les préférences individuelles ont une influence sur le choix d'établir des liens ce n'est que dans un « ensemble limité d'alternatives disponibles », c'est-à-dire un ensemble d'opportunités qui contraignent les choix possibles (p.797).

Le cadre conceptuel de la focalisation des liens met ainsi l'accent sur la nécessité de prendre en compte le contexte social en tant que « structure d'opportunité » pour expliquer l'homophilie observée dans les relations sociales. Pour cela, Feld (1982) promeut une approche structuraliste en focalisant l'étude de l'homophilie sur la structure sociale située en dehors du réseau de la relation (*extra-network social structure*) qui participe au développement des liens affinitaires. Cette structure sociale peut être appréhendée par les caractéristiques de « foyers d'activités » (*foci of activity)* autour desquels les individus organisent leurs relations sociales et leurs choix affinitaires. Selon cette approche, chaque individu est considéré en relation avec certains foyers et le lien entre deux individus dépend alors de la relation de chacun d'entre eux avec un ou plusieurs foyers d'activités communs. Les individus dont les activités sont organisées autour du même foyer tendront à se lier les uns aux autres, jusqu'à éventuellement former ce que Feld nomme un « *cluster* » (p.1016), c'est-à-dire une configuration (*pattern*) d'interactions sociales exclusives. Pour Feld, si la relation reste possible dans une rencontre en dehors d'un foyer d'activité, la plupart des relations prennent corps dans des activités organisées autour d'un foyer.

Ce cadre théorique nécessite de caractériser la nature sociologique des foyers, la distribution des relations des individus aux foyers identifiés et d'apprécier le degré avec lequel les foyers organisent, voire contraignent les interactions sociales.

Tout d'abord, comment sont définis les foyers d'activités ? Feld les définit comme des « entités sociales, psychologiques, légales, physiques autour desquelles des activités conjointes sont organisées » (Feld, 1981, p.1016 ; Feld, 1982, p.797). Ces foyers sont de nature très diverse : des individus, lieux, positions sociales, groupes et activités. Ce qui rend éligible une entité sociale à être définie comme un foyer est que celle-ci doit pouvoir potentiellement conduire ou contraindre les individus à interagir. Ces foyers d'activités peuvent être soit formels (une école), soit informels (un lieu de rencontre) et de taille plus ou moins importante (la famille, une entreprise). Les lieux de formation (écoles, universités), de travail et d'engagements (associations, syndicats) sont les foyers au sein desquels la majorité des liens se nouent (McPherson et Lovin, 1987 ; McPherson *et al.*, 2001). Ces foyers d'activités en tant que structures sociales ne sont pas inertes en ce qui concerne la création et le maintien des liens sociaux : ils contraignent les choix de liens des individus qui s'y rencontrent (Fisher *et al*. 1977 ; Feld, 1982; Stauder, 2014). Ils rassemblent des individus dont les attributs homogènes tendent à diriger le choix des relations sociales qui peuvent y être établies.

Ils diffèrent selon la catalyse ou la contrainte exercée : par exemple, la famille est un foyer non choisi par l'individu, où les interactions sont contraintes, alors que le voisinage est un foyer qui peut être choisi, où la contrainte sur les interactions sociales est faible. Certains foyers où les liens naissent et se maintiennent en son sein pour toute une vie, favorisent le soutien et la solidarité (Marsden, 1988,1990). D'autres, tels que les associations d'*almuni* des « grandes écoles » de l'enseignement supérieur peuvent avoir une influence à très long terme sur des similarités de comportements et de valeurs qui perdurent au-delà de la fréquentation de la structure au sein de laquelle les liens ont émergé (Bes, 2009, 2016).

L'ensemble des individus disponibles au sein des foyers orientera, dans une certaine mesure, les choix relationnels selon des attributs individuels particuliers (Feld, 1982). Ainsi, dans une conception structurale un foyer peut être un individu populaire c'est-à-dire une personne dont la densité relationnelle est élevée. Les relations au sein d'un réseau vont alors avoir tendance à s'organiser autour de cet individu en tant que foyer. Feld soutient l'hypothèse que la consistance de la relation entre deux individus dépend de la relation de ces individus avec un ou des foyers d'activités. Cette hypothèse est présentée comme une adaptation de la théorie de l'équilibre développée par Heider (1946, cité par Feld, 1981). Celle-ci énonce que l'intensité de l'attachement par les sentiments - d'un point de vue psychologique - entre deux individus est liée à la relation qu'entretiennent ces individus avec des entités communes. L'équilibre de la relation se réalise ainsi par la transitivité avec un objet social commun aux deux individus. Plus la relation partagée de deux individus avec les mêmes foyers d'activités est importante, plus les liens sociaux entre eux seront consistants et intenses. En effet, selon Feld (1981, 1982), la formation des réseaux de liens entre les individus et la relation de ces individus à un ou des foyers communs sont interdépendants : les relations communes avec un même foyer d'activité conduisent les individus à partager des situations mutuelles gratifiantes et des interactions valorisantes, ce qui renforce leurs liens au sein de leur réseau social. Plus le nombre de foyers partagés par les individus est important, plus la tendance à développer des liens sociaux consistants entre eux est forte.

Au-delà de la similarité des attributs individuels, le développement des liens directs entre les individus est donc déterminé en partie par les caractéristiques du contexte social au sein duquel les relations ont l'opportunité de se développer. Pour Feld (1982), les interactions sociales sont en ce sens « focalisées » par des foyers d'activités qui « tendent à produire des schémas de liens » (p.1018).

Selon le cadre théorique des relations focalisées, « les similarités d'attributs entre les individus ne conduisent pas systématiquement à des interactions focalisées et ces dernières peuvent exister en dehors des similarités de caractéristiques entre les individus » (Feld, 1982, p.1019). Les relations entre les individus n'émergent toutefois pas forcément des activités organisées autour de ces foyers et « peu importe la proportion des liens qui émergent de ces foyers, l'organisation focalisée a une signification structurale » (p. 1018).

Cette organisation focalisée des liens s'articule autour de trois facteurs structuraux majeurs : la transitivité des liens (Holland et Leinhard, 1971), les ponts structuraux (Granovetter, 1973) et la densité des relations au sein du réseau.

Le premier facteur est la transitivité. Les principales causes de la transitivité sont le nombre et le type de foyers pré-existants qui sous-tendent la relation. La transitivité s'exprime ici par la situation où « deux individus liés à un troisième peuvent partager chacun un foyer avec ce dernier et s'ils partagent le même foyer avec le troisième individu alors il est probable qu'ils soient liés tous les trois ensemble » (Feld, 1982, p.1022).

Le deuxième facteur s'exprime par les ponts structuraux. Ceux-ci sont liés au degré de partage de foyers communs par les individus. Plus les individus partagent de foyers communs, plus il est probable qu'ils soient liés par des liens consistants, c'est-à-dire forts. Inversement, si le partage de foyers communs entre des individus est faible et que ces foyers sont peu contraignants sur la relation, celle-ci sera un « lien faible » (Granovetter, 1973, 1983) qui peut potentiellement jouer le rôle de pont structural entre des groupes d'individus organisés en « *cluster* ».

Enfin, le troisième facteur est la densité. Celle-ci dépend de la transitivité et de l'existence de ponts structuraux. D'une part, si la transitivité est forte tous les individus se connaissent et la densité est très élevée. D'autre part, si tous les liens sont faibles peu d'individus se connaissent et cela génère potentiellement de nombreux ponts structuraux. La densité est alors faible. En outre, un individu impliqué dans de multiples foyers d'activité aura peu de chances d'avoir une densité de liens élevée car ses relations, issues de nombreux foyers, seront vraisemblablement peu liées entre elles, ce qui peut aboutir à une « segmentation des réseaux selon les champs d'activités » (Feld, 1982, p. 1024).

Relations sociales et relations amicales

« En vérité l'amitié (…) conforte et souvent illumine la vie de ceux qui la partagent ».
(Maisonneuve et Lamy, 1993, p. 9).

Notre étude porte sur les relations sociales affinitaires et amicales entre des individus. Nous proposons d'appréhender l'évolution de cette sociabilité affinitaire en examinant l'impact de l'un des facteurs explicatifs : le phénomène d'homophilie. Il convient toutefois, à ce stade, de définir les termes d'amitié et de relation affinitaire.

Il existe une large palette de définitions nuancées de l'amitié et des relations amicales, selon la culture des individus et la société au sein de laquelle ils vivent, comme le soulignent plusieurs travaux en ethnologie et en sociologie (Brain, 1976 ; Allan, 1979 ; cités par Degenne et Forsé, 2004). Maisonneuve et Lamy (1993) évoquent plusieurs études sociométriques comparatives des différentes conceptions et pratiques européennes et américaines de l'amitié qui soulignent que « la notion de *friendship* ne recouvre pas exactement celle d'amitié » (p. 60). Dans une étude sur l'analyse structurale des liens d'amitiés, Van de Bunt *et al.* (1999) soulignent dans la langue anglaise les différences subtiles et nuancées dans la désignation de cette relation affinitaire. Il peut ainsi s'agir, soit d'une relation du registre de l'amabilité diffuse (*friendly relationship*), soit d'une relation d'amitié (*friendship*). Les relations amiables sont plus rapides à établir que les relations amicales, moins exigeantes mais cependant moins fortes et intimes. Si elles se prêtent peu aux confidences partagées (*confiding relationship*, Marsden, 1988), les relations amiables sont toutefois une étape vers la relation d'amitié. Les « *friendly relationships* » sont ainsi des relations potentiellement amicales : elles sont en construction, en germes et elles évolueront éventuellement en relations amicales avec le temps, les circonstances et les structures au sein desquelles elles se développent, porteuses ou non d'opportunités de consolidation.

Le temps, le foyer de naissance et de croissance du lien social sont considérés comme de potentiels catalyseurs de la relation amicale. La mobilisation des ressources d'un ami pour surmonter une difficulté ou atteindre un but est le révélateur de l'effectivité de la consistance de ce lien : « *a friend in need is a friend indeed* » (proverbe anglais). Les liens amicaux offrent des opportunités d'accès à des ressources sociales, génératrices de bien-être et de satisfaction pour l'individu (Amati *et al.*, 2018). Ils peuvent également être une source d'influence et de pouvoir.

Les travaux de référence conduits par Maisonneuve et Lamy (1993) dans le domaine de la psychosociologie de l'amitié, nous éclairent sur le terme « amitié » qui croise des sens « du langage courant et d'une expérience vécue » (p. 9). Les auteurs soulignent le point suivant : « l'abus même que l'on fait du mot, plutôt que du lien, traduit implicitement une espérance ou une nostalgie ». Or, il s'agit « d'un lien privilégié, rarement éphémère (…) toujours ouvert au fil de l'existence à travers le jeu des rencontres » (p. 9).

Ce lien nécessite des rencontres, des rituels et des affinités. Tout d'abord, une nécessité de rencontre qui peut être physique et synchrone mais qui peut également se nouer à distance de façon asynchrone par un échange épistolaire ou de façon synchrone par un échange vocal. Ces rencontres nécessitent elles-mêmes des espaces, des structures d'opportunités comme nous l'avons vu précédemment, qui permettent la catalyse évoquée ci-avant. Ces espaces peuvent êtres « réels », ceux de la vie quotidienne, des études, des loisirs et du travail, ou « virtuels » comme ceux partagés *via* l'Internet par exemple. Des rituels sont également nécessaires. Nous soulignerons celui de l'échange et du partage de ressources, notamment le temps et la disponibilité qui sont par excellence les (inestimables) ressources nécessaires pour échanger avec ses pairs.

Dans notre recherche nous ne pouvons pas ambitionner de faire une étude comparative du sens et de l'expression de l'amitié selon le sexe, l'âge, l'origine ethnique ou culturelle des individus. Nous adoptons la proposition suivante de Maisonneuve et Lamy qui donne comme sens le plus courant pour les termes d'amitié et de lien d'amitié : « un ami est une personne liée à un autre par une intimité et une bienveillance mutuelle qui ne se fondent ni sur la parenté, ni sur l'attrait sexuel, ni sur les convenances sociales » (1993, p. 18). Nous considérons alors dans notre étude la relation amicale comme un lien privilégié dont l'intensité peut être hiérarchisée du niveau de l'ami, à l'ami proche jusqu'à l'ami très proche assimilable à un membre de la famille.

En ce qui concerne les affinités, bien que nous n'abordions pas les aspects psychologiques de la relation dans cette étude (cf. chapitre 1), toutefois nous précisons brièvement l'usage de ce terme dans le registre de la psychologie sociale. Pour cela, nous nous référons de nouveau aux propositions de Maisonneuve et Lamy (1993) pour qui la notion d'affinité « doit être à la fois rapprochée et distinguée » des notions de préférence (dimension individuelle) et d'affiliation (dimension collective) (p. 16).

D’une part, la préférence peut avoir une dimension bilatérale ou unilatérale tandis que l’affinité est bilatérale même si le lien affinitaire partagé n’a pas la même intensité réciproque. D’autre part, l’affiliation réfère à l’association d’un individu avec un ou des autres ou l’intégration dans un groupe, tandis que l’affinité implique une relation dyadique.

Cependant, nous nous éloignerons des choix de Maisonneuve et Lamy (1993) en ce qui concerne la caractérisation de l’affinité. Dans le cadre de leurs travaux, ces auteurs choisissent de caractériser l’affinité comme dilective plutôt que sélective ou élective. Ainsi, ces derniers semblent exclure de la caractérisation de la relation affinitaire les rapports instrumentaux et intéressés. Dans le cadre de notre investigation, nous estimons que le « choix vécu affectif gratifiant pour ses agents » (Maisonneuve et Lamy, 1993, p. 17) n’exclut pas une démarche guidée par un intérêt pour accéder à des ressources du domaine affectif (confidence, soutien) ou instrumental (service, aide). D’une part, nous considérons que la relation amicale constitue une ressource de nature sociale non monétaire qui permet d’exprimer des sentiments, partager des confidences de bénéficier d’une aide en cas de besoin. Ceci nous apparaît en accord avec Maisonneuve et Lamy (1993, p. 18) qui expriment que « ces besoins socio-affectifs tendent à se satisfaire à travers les relations personnelles que les sujets ont entre eux ; deviennent amis ceux pour qui ces relations sont source d’une satisfaction mutuelle ». D’autre part, nous considérons qu’individu qui dispose d’un ensemble de liens amicaux peut bénéficier de ressources diverses - aide, service, conseils - potentiellement mises à disposition par ses relations pour faciliter l’atteinte d’un but ou dans le cadre d’un échange réciproque.

Ainsi, les relations amicales, tout en étant libres, généreuses et gratuites, ne sont dépourvues ni de déterminants sociaux ni d’attentes de gratifications ou de bénéfices autres que ceux qui relèvent uniquement du registre affectif. Or, pour évoluer de la rencontre aimable (*meeting)* à l’appariement amical (*mating)* (Verbrugge, 1977 ; Van Dujin, 2003) ces relations doivent être nourries par un investissement relationnel, des échanges intenses et entretenus selon des modalités variées afin de demeurer dans le registre de l’amitié et ne pas décliner vers celui de l’amabilité, moins gratifiant. Cet investissement représente un coût en temps et en implication qui peuvent être mis en relation avec des bénéfices potentiels. Dans notre étude nous intégrons donc les dimensions de coût en termes d’investissement relationnel et de bénéfices en termes de ressources sociales. Ces coûts et ces bénéfices sont considérés comme des composantes de la dynamique d’une relation.

Pour cela, nous adoptons d’une part, le cadre théorique du choix rationnel et, d’autre part, le concept de capital social, présentés dans la section suivante.

L'individu acteur et l'investissement relationnel

Dans le cadre de notre analyse structurale, nous considérons que l'investissement relationnel nécessaire pour faire évoluer une relation du « *meeting* » au « *mating* » représente un coût. En outre, nous considérons que selon le coût et l'objectif poursuivi les individus, en tant qu'acteurs de la relation, adoptent un comportement rationnel dans les choix électifs et sélectifs de leurs liens sociaux. Granovetter (1985) invite à prendre en considération ce postulat de « l'action rationnelle » en soulignant que « bien que l'assomption de l'action rationnelle soit toujours problématique, c'est une bonne hypothèse de travail qui ne devrait pas être abandonnée facilement » (p.506). Par conséquent, nous convoquons la théorie du choix rationnel selon laquelle l'acteur compare les avantages (bénéfices) et les inconvénients (coûts) d'établir une relation et qu'il agit selon son intérêt en choisissant des liens sociaux qui disposent des ressources dont il a besoin pour atteindre ses objectifs personnels ou professionnels (Coleman, 1988, 1990). Dans le cadre de l'analyse structurale une des hypothèses majeures est que, d'une part, les choix des individus participent à former la structure des liens sociaux et que, d'autre part, la structure agit également sur les choix des individus dans la sélection de leurs relations sociales. Nous adoptons plus spécifiquement le modèle de rationalité « élargie » énoncé par Forsé et Parodi (2004). Selon ce modèle, l'acteur n'agit pas exclusivement selon son seul intérêt. En effet, il peut réduire ses attentes et procéder de façon non optimale tout en œuvrant à faire évoluer la structure des relations vers l'atteinte de ses objectifs. L'acteur agit ainsi de façon « rationnelle et raisonnable » (Degenne et Forsé, 2004; Forsé, 2008).

L'individu « acteur » étant ainsi caractérisé, il reste à cerner quels peuvent être les ressorts de son action, c'est-à-dire le moteur de la dynamique de ses relations sociales. Dans la section suivante, nous présentons comme proposition de moteur de l'action les ressources du capital social contenu dans le réseau constitué par les individus. Nous explorons comment les relations sociales permettent à l'acteur de satisfaire certains de ses besoins grâce aux liens qu'il noue avec d'autres au sein du réseau social, selon sa position dans le réseau et les objectifs qu'il poursuit.

Des bénéfices asynchrones

Nous considérons que les acteurs construisent des relations selon un but et qu'ils les entretiennent tant que celles-ci continuent à produire des bénéfices. Par conséquent, les ressources générées peuvent se déprécier en tant que ressources.

En effet, si le lien relationnel n'est pas suffisamment entretenu entre les individus et qu'il décline en intensité ou qu'il s'estompe, l'obligation morale de réciprocité est susceptible de s'estomper au fil du temps. Le remboursement de la ressource mise à disposition est moins certain et la ressource générée chez le créditeur peut finalement devenir un mauvais placement : « *placement of trust has been unwise (...) bad debts that will not be repaid* » (Coleman, 1998, p.102). Dans ce sens, Boissevain (1974 ; cité par Flap, 2002) souligne que l'importance d'une relation est déterminée par des investissements opérés dans le passé et dans une attente d'aide dans le futur en réponse à cet investissement. Ainsi, « les individus investiront et désinvestiront dans du capital social selon leur attente d'une aide future » (Flap, 2002, p. 37).

« Investing in other is similar to what game theory calls playing a trust game ».
(Flap, 2002, p. 38).

La question se pose alors du réel bénéficiaire du capital social, par exemple dans le cas d'une demande de service : est-ce celui qui demande un service ou celui qui a répondu à la demande et qui détient alors la « dette » ? Pour Coleman (1988), la mobilisation d'une ressource détenue par autrui augmente le « fond » de capital social disponible de celui qui détient la ressource mobilisée (p.117). Il est alors envisageable de considérer qu'il y a au moins deux bénéficiaires dans chaque action de mobilisation du capital social. Premièrement, l'individu qui sollicite des ressources d'autrui au temps i : il tire un bénéfice immédiat de la mobilisation de la ressource « empruntée » à autrui. Deuxièmement, l'individu qui fournit cette ressource au temps t_i : selon le principe de réciprocité, il peut espérer stratégiquement « valoriser » son action envers le demandeur pour une sollicitation au temps t_{i+n} si un besoin émerge pour l'atteinte de ses propres buts. Flap (2002) souligne que ceci constitue une des hypothèses qui se situent au cœur de la théorie du capital social. Il fait ainsi référence aux propositions de Coleman (1988, 1990) en soulignant que « l'importance du capital social, dans la vie sociale, est très liée au fait d'être un cas d'aide non simultanée (*delayed or generalized reciprocity*), ce qui est normalement, au temps donné, considérablement plus valorisable par le récipiendaire que ce n'est couteux au donneur » (Flap, 2002, p.37).

L'importance du temps

« The future can therefore cast a shadow back upon the present and thereby affect the curent strategic situation » (Axelrod, 1984, p.12).

Le temps intervient donc de façon centrale dans les conditions d'exploitation du capital social au sein du réseau social d'un individu. Dans son ouvrage « The evolution of cooperation », Axelrod (1984) défend l'idée que le temps agit sur le taux de dépréciation de l'aide attendue dans le futur. Prenant appui sur le « dilemme du prisonnier », qui illustre les positions de conflit ou de coopération entre deux individus qui ont à choisir entre l'intérêt collectif et l'intérêt individuel, Axelrod explique que :

> *« Ce qui rend possible l'émergence de la coopération, c'est le fait que les joueurs aient la possibilité de se rencontrer de nouveau. Cette possibilité signifie que les choix faits aujourd'hui ne déterminent pas seulement les résultats courant de l'action, mais peuvent aussi influencer les choix futurs du joueur » (Axelrod, 1984, p.12).*

Ainsi, l'évolution du temps peut jeter l'ombre du futur (*shadow of the futur*) sur le présent et influencer l'acteur dans le moment de l'action. Ce risque potentiel augmente le coût de l'investissement relationnel de l'individu (*the discount rate*) pour entretenir au long court la relation (au moins tant qu'il n'y a pas eu de retour). Le délai entre l'investissement et le retour sur l'investissement augmente le risque pour l'individu et par conséquent l'importance de la confiance dans la relation sociale, en tant qu'élément constitutif du capital social. L'attente du donneur est que le bénéficiaire immédiat n'agisse pas de façon opportuniste et selon ses seuls intérêts, mais qu'il agisse de façon rationnelle et raisonnable, comme nous l'avons vu précédemment. Ainsi, les coûts et les bénéfices issus de l'entretien d'un lien social, les investissements relationnels, l'ombre du futur, le temps et les alternatives relationnelles offertes par les liens croisés sont des éléments qui jouent sur la décision de l'acteur « d'investir ou de désinvestir » un lien avec autrui (Flap, 2002).

Un moteur de l'action : les ressources du capital social

« (...) no engine for action (...) no internal spring of action » (Coleman, 1988, p.96).

Le concept de « capital social » alimente une littérature riche et diverse depuis une trentaine d'années avec des développements centrés soit sur l'individu, soit sur la société selon une considération de bien public (Putman, 1993, 1995). Dans cette recherche, nous convoquons les travaux fondateurs de Coleman (1988) ainsi que ceux de Granovetter (1973), Burt (1992, 2000) et Lin (1982, 2005) relatifs aux concept de capital social et d'encastrement des ressources de ce capital dans les structures sociales.

Coleman (1988) propose trois principales formes qui structurent le capital social.

La première est relative aux obligations observées au sein d'une communauté. Pour être respectées, ces obligations nécessitent un niveau de confiance entre les individus qui forment la communauté. La confiance au sein de la communauté dans le respect mutuel des obligations est ainsi une des premières dimensions du capital présent dans la communauté.

La deuxième forme découle des obligations : elle est constituée de l'ensemble des normes et des sanctions présentes dans la régulation sociale d'une communauté. Celles-ci influencent les comportements des individus et facilitent le contrôle social.

La troisième forme de capital social proposée par Coleman (1988) est relative à l'information.

Dans le cadre de notre étude, c'est cette troisième forme qui retient le plus notre attention. En effet, au sein d'une communauté de jeunes professionnels d'ethnies et de provenances géographiques variées, en reprise d'études au sein d'une des plus grandes universités mondiales, l'information peut revêtir une valeur cruciale en tant que ressource pour accéder à des conseils et à de futurs contacts professionnels, voire des opportunités d'emploi.

Avant de déplier le concept de capital social et ses déclinaisons, il convient préalablement d'évoquer succinctement les critiques qui lui sont adressées.

« Social capital constitues an unanalysed concept, it signals to the analyst (...) that something of value has been produced for those actors who have this resource available and that the values depends on social organization » (Coleman, 1988, p.101).

La définition proposée par Coleman (1988) peut être perçue insatisfaisante car elle rassemble tout à la fois « un concept, un outil, une fonction et une ressource » (Portes, 1998, Durlauf 1999, cités par Ponthieu, 2006) et les raisonnements relatifs à ce concept peuvent sembler « circulaires et tautologiques » (Lévesque et White, 1999, p.25) car ils ne distinguent pas suffisamment ce qui crée le capital social et ce que produit le capital social. Par conséquent, cette conception se risque au « piège de l'induction factice de la présence du capital social en présence de ses effets supposés » (Lévesque et White, 1999, p.27).

Pour Coleman (1988), le concept de capital social demeure « non analysé » (*unanalyzed concept*) et nécessite d'être déplié (*unpack the concept*) afin de « clarifier les composantes de l'organisation sociale qui contribuent à la valeur produite » (p.101). Toutefois, s'ils nécessitent d'être affinés, plusieurs éléments de cette approche conceptuelle sont fertiles pour notre d'étude.

Plusieurs des auteurs mentionnés ci-avant expriment l'idée que les ressources sociales constituent le capital social qui est différent du capital humain, économique et culturel. Les individus investissent dans ce capital social leur travail relationnel et peuvent en espérer des bénéfices pour atteindre leurs buts. Dans la partie suivante, nous nous intéressons aux dépliages du concept de capital social, aux mécanismes structuraux qui le génèrent et qui peuvent contribuer, dans le cadre de notre étude, à expliquer comment et pourquoi les relations au sein d'un réseau social se forment et se déforment. Ainsi, nous soulignons quelques propositions saillantes d'auteurs qui positionnent le capital social comme une ressource inhérente aux pratiques de sociabilité. En particulier, nous valorisons l'approche de Lin (1982, 2005) dans sa réflexion sur la théorie des ressources sociales.

Par le concept de capital social, Coleman (1988) développe un outillage théorique qui vise à intégrer différentes approches dans la description de l'action sociale qui opposent des sociologues et économistes américains dans les années 1980, « période où les économistes et les sociologues engageaient un dialogue un peu étrange, fait d'incursions de chacun dans le territoire de l'autre, d'escarmouches à l'occasion d'un débat dans une revue, d'emprunts réciproques et de rejets virulents » (Grossetti, 2015, p.1).

La première approche, portée par des sociologues, considère que l'acteur socialisé est guidé dans son action par un contexte constitué de normes sociales, de règles et d'obligations qu'il internalise et qui contraignent ses choix. L'acteur est alors considéré comme le seul produit de son environnement social. La critique portée à ce courant de pensée est que, dans ce cadre, il est difficile de discerner le « moteur et le ressort de l'action » de l'acteur qui lui donnent une direction pour agir (Coleman, 1988, p.96). Cette critique rejoint le questionnement de Wrong sur la conception dans la sociologie fonctionnaliste d'un homme « sur-socialisé » et d'une société « sur-intégrée » (Wrong, 1961, p.184). Pour Wrong, cette conception répond de façon insatisfaisante à la question fondamentale de Thomas Hobbes (1651) soulevée dans le Léviathan : comment l'homme devient-il conciliant avec les contrôles sociaux ? La réponse apportée par la sociologie se résumerait, selon Wrong, à considérer que l'homme intériorise les normes sociales qui sont non perçues comme des poids et qu'il agit essentiellement pour promouvoir une image favorable de lui-même, c'est-à-dire conforme aux attentes des autres. Cette conception dénierait la possibilité pour l'individu d'être autre chose qu'une personne complétement socialisée. Sans verser dans « un excès de psychologisme », l'auteur invite à intégrer le point de vue selon lequel l'homme est « un être social bien que non complétement socialisé » (Wrong, 1961, p.192).

La seconde approche que Coleman souhaite intégrer à l'aide du concept de capital social est celle portée par des économistes. Parmi ces derniers, certains considèrent que l'acteur est guidé par des buts et qu'il agit en conséquence selon son seul intérêt et de façon indépendante afin de maximiser l'utilité de son action. Selon cette approche, l'acteur agit comme un individu « sous-socialisé ». Pour Coleman (1988), ce postulat va à contre-courant de la réalité empirique en écartant l'importance des contraintes exercées par l'environnement social sur l'individu ainsi que le rôle des normes, de la confiance et des organisations sociales dans le fonctionnement de la société. D'une part, Coleman critique les approches individualistes de certains économistes qui considèrent l'économie comme constituant une sphère qui serait distincte dans la société, où les transactions ne seraient plus déterminées par les obligations sociales et affinitaires mais seulement par les calculs rationnels du gain individuel. D'autre part, Coleman critique également l'approche strictement holiste qui promeut que les comportements individuels peuvent être complétement déduits de la structure sociale. Pour l'auteur, l'action de l'individu est toujours socialement située et ne peut pas être expliquée exclusivement par la poursuite de ses intérêts individuels ou par les contraintes exercées par la structure sociale.

Le capital social est ainsi considéré comme un « outil conceptuel » qui permet d'intégrer ces deux traditions intellectuelles sans verser dans le « pastiche » (Coleman 1988, p.96, 97). Sa valeur est « d'identifier certains aspects de la structure sociale par leurs fonctions » (p.101). L'auteur souligne que la fonction identifiée par le concept de capital social est la valeur pour l'acteur des aspects des structures sociales qu'il peut utiliser comme des ressources pour atteindre ses buts. Ainsi, ce capital est constitué d'un ensemble « d'entités » (*entities*) qui partagent deux caractéristiques : « elles relèvent d'un aspect de la structure sociale et elles facilitent les actions de l'individu au sein de la structure » (p.98). Ce capital serait productif comme d'autre formes de capitaux (humains, physiques). Il rendrait possible l'atteinte par les acteurs de leurs objectifs grâce à l'utilisation des ressources présentes au sein de leur réseau social.

La question se pose alors du lieu où se situe ce capital que l'individu peut mobiliser via ses relations sociales.

Pour Coleman (1990) le capital social est inhérent à la structure formée par les liens entre les individus. Il n'est ni englobé (*embodied*) dans un équipement productif ni généré par la transformation de matériaux en outils productifs, comme pour le capital physique, ni en « en dépôt » (*lodged*) chez les acteurs suite à un accroissement de leurs compétences, comme pour le capital humain. L'auteur situe ce capital dans les lignes qui relient les points formés par le capital humain, c'est-à-dire les liens qui relient les individus - les nœuds - d'un réseau social. Lin (2005) abonde dans ce sens en argumentant que le capital social est enraciné (*rooted*) dans les réseaux sociaux et les relations sociales. L'hypothèse centrale est que les individus utilisent leurs ressources personnelles et empruntent celles des autres, disponibles au sein de leur réseau social. Par conséquent, la valeur du capital social dépend de la structure sociale considérée et des changements relationnels.

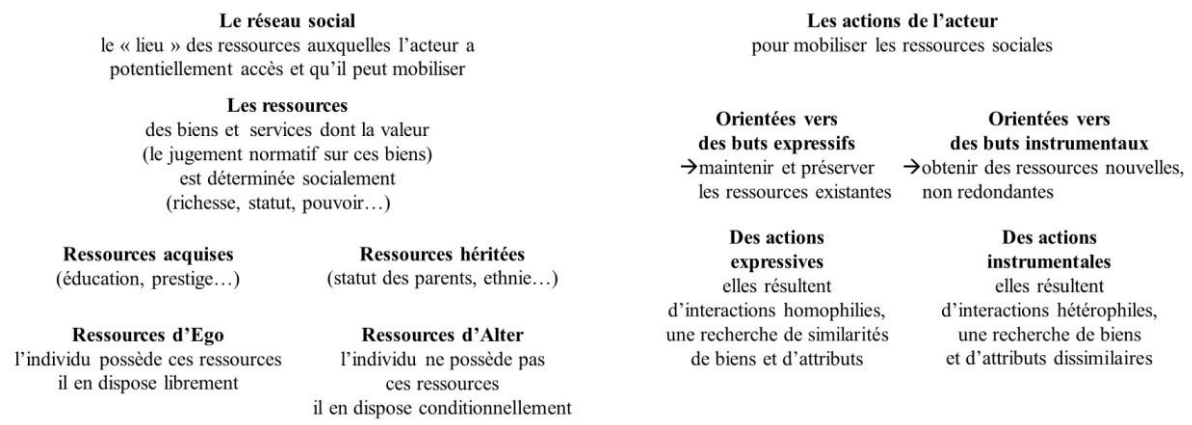
Les ressources sociales

« Les réseaux figureraient un peu comme les 'lieux' des ressources et les contacts comme l'accès à ces lieux » (Ponthieu, 2016, p.28).

Nous avons évoqué que Coleman (1988) considère que les relations sociales tissées entre des individus d'une communauté peuvent être appréhendées comme un investissement de ressources au sein de cette structure sociale. Le capital social est ainsi considéré comme l'ensemble des « ressources qu'une communauté est susceptible de procurer à ses membres » (Forsé, 2002, p. 289). Lin (1995) propose un cadre de réflexion qui permet d'aller un peu plus loin dans la définition et l'opérationnalisation du capital social. Selon l'auteur, les ressources sociales se situent au centre du capital social. Nous exploitons cette approche dans notre étude.

Lin (1982, 1995, 2005) développe la définition du capital social en prenant comme cadre de référence la théorie sociologique des ressources. Dans ce cadre théorique, les actions individuelles sont décrites dans un contexte structural. Il s'agit de considérer que chaque acteur dispose de certaines ressources personnelles et qu'il peut, à travers ses échanges avec d'autres au sein de son réseau social, accéder aux ressources d'autrui et mobiliser celles dont il a besoin pour atteindre ses buts (Lin, 1995). L'individu n'en dispose pas librement mais elles peuvent être empruntées ou capturées selon des conditions, dont la principale est la réciprocité. Nous proposons de synthétiser les propositions de Lin sur la figure n°1 ci-après.

Figure n°1 : réseau, ressources, acteur, buts et actions.
Composition élaborée d'après Lin (1982, 1995, 1999, 2001, 2005).



Selon les ressources recherchées par l'acteur, les actions entreprises par celui-ci peuvent être catégorisées en deux groupes selon leur caractère plutôt expressif (je recherche un ami auprès duquel je peux me confier) ou plutôt instrumental (je recherche un ami qui peut m'aider) (Lin, 1995 ; Degenne et Forsé, 2004). Nous considérons ces ressources comme étant plutôt expressives ou plutôt instrumentales, sans considérer ces deux ensembles étanches ou disjoints. Ces ressources sont de nature très variée : il peut s'agir par exemple d'un carnet d'adresse de contacts susceptibles de faciliter une recherche d'emploi, de bénéficier d'influence, de conseils ou de confiance, de solliciter un service tel qu'un prêt d'argent, une garde d'enfants ou une aide en cas de maladie (ressources instrumentales). Il peut également s'agir de pouvoir échanger des sentiments ou des interrogations, de solliciter de la compréhension, de l'empathie, de l'approbation ou du soutien moral, qui sont des constituants essentiels du bien-être de l'individu (ressources expressives).

Afin de préciser sa conceptualisation du capital social, Lin (2005) distingue trois « couches » (*layers*) comme lieux des ressources.

Tout d'abord, une couche intérieure (*inner layer*). Cette couche est susceptible d'enfermer (*confining)* les individus que de faciliter un mouvement, une mobilité sociale. Les ressources de ce lieu sont équivalentes entre l'individu et ses contacts. Cette couche est caractérisée par des ressources expressives, telles que le partage de sentiments, le soutien mutuel, accessibles et mobilisables par des liens forts et réciproques entre des individus similaires qui partagent de nombreux intérêts.

Ensuite, Lin (2005) distingue une couche intermédiaire (*intermediary layer*). Elle est le lieu de ressources nouvelles pour l'individu, d'une nature plutôt instrumentale. Cette couche est caractérisée par un mélange de relations proches et distendues qui relient inégalement les individus qui partagent certains intérêts, avec une circulation des ressources qui n'est pas systématiquement réciproque entre les individus.

Enfin, l'auteur distingue une couche extérieure (*outer layer*). Celle-ci forme un espace où les individus possèdent une forme d'identité commune par l'affiliation à une institution (politique, religieuse, associative ou universitaire). Ces individus peuvent ou non interagir entre eux. Leurs relations sont médiatisées par l'institution qui leur fournit « le fond de l'appartenance partagée » (*belongingness*) (Lin, 2005, p. 12).

Dans le cadre de notre étude, nous relevons que si l'acteur éprouve la nécessité de plus de ressources et de plus grande valeur pour accomplir ses buts ; il pourra essayer d'atteindre une couche en dehors de la sienne à l'aide de ponts sociaux. Ces ponts peuvent être des relations distendues, nommées par Granovetter (1973) des « liens faibles ». Ces liens faibles peuvent permettre à l'acteur d'accéder à un lieu du réseau social où circulent des ressources plus hétérogènes que les siennes, à condition d'« enjamber » des « trous structuraux » (Burt, 1992). Cela correspond à une stratégie d'action instrumentale. Toutefois, si l'individu n'a pas besoin de nouvelles ressources, le bénéfice accessible dans sa « couche » est celui produit par la similarité, les relations denses et réciproques. Par conséquent, il apparaît que les actions instrumentales génèrent une évolution structurale du réseau favorable à l'ouverture et aux changements, alors que les actions expressives œuvrent plus vers un état de stagnation (*stasis*), de stabilité et d'équilibre, favorable à la cohésion et à la fermeture (Burt, 2000, p.355).

En développant les concepts « d'encastrement » et « d'enchâssement » (*embeddedness*), Granovetter (1973) propose de considérer l'action rationnelle de l'individu comme encastrée dans les relations sociales (*the embeddedness of economic behavior*, p.842). Le terme d'encastrement évoque le fait que les réseaux formés par les liens « sont pour les individus des ressources autant que des contraintes. Pour Granovetter (1973), l'encastrement explique par exemple le rôle des relations personnelles et des réseaux sociaux dans la production de la confiance, de la réputation et dans la réduction de la malversation. L'encastrement n'est ni une dissolution ni un déterminisme, il relève plutôt d'une « dépendance » (Grossetti, 2015, p.3).

A partir de l'exemple d'une transaction marchande où la réputation et la loyauté des acteurs jouent un rôle clef, l'auteur argumente que l'appréciation de la fiabilité d'un individu est meilleure si l'on prend en compte l'information donnée sur cet individu par un informateur de confiance (*a trusted informant*, p.490) qui a déjà commercé avec ce dernier, plutôt que d'estimer le ratio du coût des dommages causés sur la réputation de l'individu s'il agissait déloyalement et du bénéfice de l'opportunité d'agir ainsi selon un point de vue strictement économique.

Par ailleurs, Granovetter rejoint l'intérêt que Coleman porte à la notion de fermeture (*closure*) au sein d'un réseau social : la réputation, élément constitutif de la confiance interpersonnelle, ne peut pas se construire dans des structures ouvertes mais seulement dans des structures fermées dans lesquelles chaque individu est connecté à tous et où aucun ne peut se soustraire au regard évaluatif de l'autre (Coleman, 1988, 1990). En particulier, Granovetter (1973) développe la notion de force du lien social. La force d'un lien social est, selon l'auteur, une combinaison de plusieurs éléments : la quantité de temps accordée à la relation, l'intensité émotionnelle de la relation, la confiance mutuelle entre les individus reliés et l'existence de services rendus de façon réciproque entre les individus impliqués dans cette relation. Ces éléments constitutifs du lien seraient corrélés mais toutefois relativement indépendants (p.1361). L'auteur souligne la différence d'apports des liens forts et faibles pour l'acteur dans l'accomplissement d'un but tel que la recherche d'un emploi par exemple (Granovetter, 1974 ; Forsé, 1997, 1999). Celui-ci distingue au sein des réseaux de relations sociales, d'une part, des liens sociaux forts formés à partir de relations intenses et entretenues avec des amis et, d'autre part, les liens sociaux faibles formés par des relations plus épisodiques et distendues avec des connaissances. Les liens faibles peuvent s'avérer plus bénéfiques que les liens forts car ils apportent des ressources nouvelles, moins redondantes que celles fournies par ces derniers.

Les termes et les concepts évoqués précédemment - cohésion, fermeture, liens faibles, trous structuraux - nous conduisent à expliciter plus précisément l'articulation entre le capital social, l'acteur rationnel, les ressources sociales et l'analyse structurale qui est le cadre dans lequel nous positionnons notre étude.

« *All social relations and social structures facilitate some forms of social capital* »
(Coleman, 1988, p.105).

Deux caractéristiques structurales apparaissent essentielles pour la production du capital : la réciprocité et la fermeture structurale (*closure*). Une caractéristique structurale crée l'opportunité de son utilisation : l'ouverture structurale.

La réciprocité s'illustre assez simplement dans les relations sociales. Par exemple, un service rendu est une ressource pour l'individu qui en a besoin. Selon le principe de réciprocité des relations, le service rendu crée une dette - que Coleman (1988) nomme « une lettre de créance », *credit slip* - c'est-à-dire une ressource pour une éventuelle utilisation dans le futur. Cette dette est détenue par l'individu qui a rendu le service. Cette conception fait référence aux théories sur le don (Mauss, 1923)[3] : l'individu cumule des ressources en donnant des ressources ce qui crée une dette chez le récipiendaire. Le capital social est ainsi constitué de l'ensemble des bénéfices attendus dans le futur d'un investissement dans les relations sociales au sein du réseau.

La notion de fermeture structurale (*closure)* est un élément qui joue dans le partage des normes, des obligations et des attentes. En effet, la fermeture produit une forte cohésion qui favorise la confiance partagée, la fiabilité (*trustworthiness*). Elle facilite l'accès à l'information pour les individus du réseau. Ainsi, par exemple, lorsque les membres d'une communauté de parents sont tous en relation dans une configuration de fermeture (par exemple un réseau complet connexe) cela assure la diffusion des mêmes normes, récompenses et sanctions liées au comportement des enfants (Coleman, 1988, p.107-108). Dans ce cas, la fermeture (*closure*) offre les avantages de la cohésion et procure les bénéfices du contrôle normatif. Ce type de bénéfice peut être dégagé des ressources offertes par des personnes similaires (« *like-minded people* », Burt, 2000, p.356) ou qui occupent des positions structurales équivalentes. Cela est à rapprocher du principe d'homophilie et de la recherche de ressources expressives (Lin, 1995) que nous avons vu précédemment.

---

[3] Mauss, M. (1923). Essai sur le don. Forme et raison de l'échange dans les sociétés archaïques. *L'Année Sociologique*, seconde série, 1923-1924.

Toutefois, les bénéfices issus de la fermeture et de l'équivalence structurale peuvent être redondants. En effet, les individus fortement connectés entre eux sont susceptibles de disposer des mêmes ressources homogènes selon le principe d'homophilie évoqué précédemment et les relations structurellement équivalentes sont susceptibles de procurer à l'acteur des ressources redondantes.

Ceci nous invite à considérer le concept d'ouverture structurale proposé par Burt (1992). Celui-ci nous apporte une formalisation intéressante pour comprendre comment l'acteur peut trouver dans certaines configurations structurales des bénéfices nouveaux, hétérogènes. A l'aide de la métaphore précédemment mentionnée de « trou structural » (*structural hole*), Burt (1992) définit la configuration d'un individu positionné dans le réseau telle que les acteurs avec lesquels il entretient des relations ne sont pas eux-mêmes directement en relation. Les trous structuraux caractérisent ainsi l'absence de relations entre des pairs structuralement redondants. Ils contribuent à séparer les ressources non redondantes, c'est-à-dire les informations qui sont « plus additives que superposées » (Burt, 2000, p.353). Dans une configuration où le réseau social d'un individu contient des trous structuraux, s'il est en position de les franchir par des relations « passerelles » alors son capital social peut s'enrichir de ressources nouvelles. Selon la position de l'acteur dans le réseau social, l'ouverture créée par des trous structuraux peut être une source d'opportunité d'accès à des ressources recherchées, plus rares que dans la configuration fermée. Pour Burt (2000), le capital social est une métaphore de l'avantage dont peut bénéficier l'acteur dans un marché ou les individus échangent des biens, des services et des idées dans la poursuite de leurs intérêts. L'auteur considère que le réseau et la position de l'acteur au sein de celui-ci constituent le capital social, c'est à dire l'ensemble des ressources qu'il peut mobiliser. Le réseau est considéré comme porteur de ressources mais également de contraintes pour l'acteur. Toutefois, ces contraintes pèsent d'autant moins sur ses choix, guidés par l'atteinte d'objectifs, qu'il n'est pas dans une situation de fermeture (*closure*) dans ses relations avec ses pairs : en effet, ces derniers ne peuvent pas le contraindre ensemble.

Cette configuration structurale est susceptible d'offrir à l'individu des opportunités et des avantages. En effet, en exploitant des possibilités de stratégies profitables ce dernier peut se positionner comme un nœud central et jouer le rôle d'intermédiaire *(gatekeeper*) ou de négociateur (*broker)* entre ses relations déconnectées. Il peut contrôler la diffusion de l'information, la filtrer à son avantage et saisir les opportunités générées par l'absence de lien entre ses relations. Ainsi, pour Burt (2000) : « les trous structuraux créent un avantage compétitif pour l'acteur dont les relations enjambent ces trous » (p. 353), car il est placé dans une position d'intermédiation.

Un des points importants que nous soulignons ici est l'avantage putatif donné à l'acteur, pour accéder à des ressources nouvelles, lié à la diversité des individus avec lesquels il choisit de s'appairer. Cet avantage est potentiellement plus intéressant que leur nombre de pairs avec qui il noue des liens. Par conséquent, d'une part, nous considérons la diversité des individus du point de vu de l'hétérogénéité de leurs attributs socio-démographiques et de leurs positions structurales. D'autre part, dans le cadre du suivi d'une communauté de jeunes professionnels en reprise d'études dans une université de renommée mondiale, nous soutenons que pour ceux-ci l'appariement hétérogène, c'est-à-dire l'établissement de relations avec des pairs de d'attributs socio-démographiques et degrés de connectivité différents, est une source de bénéfice qualitativement plus importante que celle fournie par l'appariement homogène.

Pour notre étude, nous utilisons donc plusieurs concepts relatifs aux réseaux sociaux et aux ressources sociales proposés par Coleman (1988), Lin (1982, 1995, 2005), Burt (1992, 1995, 2001) et Granovetter (1973, 1983). Nous adoptons la définition du capital social comme « l'investissement d'un individu dans ses relations avec d'autres » et qui ne peut exister que « si l'on examine les relations d'un individu et les ressources accessibles grâce à ces relations » (Lin, 1995, p.701). Nous considérons que les ressources sociales constitutives de ce capital sont un des moteurs de l'action de l'acteur. Elles participent à expliquer au cours du temps son investissement/désinvestissement dans l'établissement de liens au sein de son réseau. A travers le réseau des relations sociales nous considérons que, selon ses buts, l'acteur accède et mobilise par son travail relationnel des ressources détenues par des pairs susceptibles de générer un bénéfice en retour de son investissement (Lin, 2005).

L'homophilie, nous l'avons vu, facilite l'établissement rapide de relations entre des individus qui partagent des attributs sociaux et des modes vies similaires, ce qui leur permet ainsi de mobiliser des ressources de leurs semblables. Toutefois, les ressources sociales détenues par des individus semblables tendent à être homogènes et redondantes d'une relation homophile à une autre. L'homogénéité sociale générée par l'homophilie dans les relations sociales peut alors se révéler comme un frein pour mobiliser des ressources sociales hétérogènes, moins redondantes et potentiellement de plus grande valeur car moins partagée (Mc Pherson *et al.*, 2001). Les liens sociaux hétérophiles permettent d'accéder à des ressources variées, plus rares, potentiellement meilleures ou plus valorisées que celles partagées dans le collectif homophile. L'action instrumentale nécessaire pour les mobiliser nécessite l'accès à des individus qui occupent des positions différentes (Lin, 1995, p.689). Cette action instrumentale peut être facilitée par des liens faibles qui serviront de ponts, c'est à dire d'intermédiaires (*gatekeepers*) (Granovetter, 1973).

Ainsi, nous définissons l'action de l'acteur comme la mobilisation soit de ressources homogènes, par l'accès direct ou indirect à des individus similaires, soit de ressources hétérogènes, par l'accès direct ou indirect à des individus dissimilaires.

Enfin, nous soulignons que pour Coleman (1988) le caractère multiplexe des relations sociales facilite la circulation des ressources d'un réseau à l'autre. En effet, les ressources issues d'une relation existante au sein d'un réseau peuvent être utilisées dans d'autres réseaux de relations avec, toutefois, une possible dépréciation de sa valeur hors du contexte de sa création. La prise en compte de la multiplexité des relations apparaît donc importante pour appréhender la dynamique d'évolution des liens entre les individus lorsque celle-ci est expliquée par la recherche de ressources sociales. Nous précisons l'intérêt de considérer l'entrelacement des relations sociales dans la dernière partie de ce chapitre.

L'entrelacement des relations sociales : un récit multiplexe de la sociabilité

« Les relations et donc les réseaux sont des histoires, il faut donc les analyser comme tels : la prise en compte du multi-niveau, cercles et relations, est importante ». Degenne (2013, p.38).

Malgré l'important développement des supports numériques de communication synchrone et asynchrone entre les individus, l'observation des interactions en face à face demeure un mode de communication essentiel pour qui souhaite étudier les comportements humains au sein d'une communauté. Cette modalité de communication est communément considérée comme le prototype de l'interaction sociale. Toutefois, elle semble perdre sa centralité dans l'appréhension du développement des liens de sociabilité entre les individus (Mastrandrea *et al.*, 2015). En effet, elle partage de plus en plus sa place dans le développement des liens de sociabilité avec d'autres médiations interactionnelles - vocales et textuelles - qui sont proposées par les outils de communication à distance tels que le téléphone, la messagerie électronique et l'Internet.

Néanmoins, si ces modalités de communication en nombre croissant prennent une part de plus en plus grande dans la construction de la sociabilité, elles n'apparaissent pas aujourd'hui se substituer l'une à l'autre et ne créent pas un espace hermétiquement séparé avec l'univers relationnel en co-présence physique (Cardon et Smoreder, 2014 ; Bailliencourt, 2007). Ainsi, plusieurs recherches montrent que la multiplication des modalités de communication ne modifie pas considérablement les formes de sociabilité. Par exemple, les observations issues du projet « Entrelacs », soulignent que les formes globales de sociabilité sont plus dépendantes de l'évolution du cycle de vie des individus que de la multiplication des usages des médias disponibles (Smoreda *et al.*, 2007).

Smoreda *et al.* (2007) observent d'avantage l'émergence d'usages « entrelacés » qu'une substitution d'une modalité de communication par une autre. En particulier, il apparaît que plus l'individu fait appel à une grande variété de modalités d'interactions, plus celui-ci donne *in fine* de l'importance au face-à-face où l'expérience de l'autre est la plus complète : « *The most important experience of others takes place in the face-to-face situation, which is the prototypical case of social interaction. All other cases are derivatives of it* » (Berger et Luckman, 1966, p. 43). Ainsi, « plus on se sent proche de quelqu'un, plus on convoque une grande variété de dispositifs d'interactions » parmi ceux qui sont partagés (Bailliencourt *et al.*, 2007, p.112).

Toutes ces modalités de communication, en face-à-face, par téléphone ou par « épistole électronique » ne dessinent donc pas des univers de sociabilité très différents : ils s'entrelacent étroitement (Cardon et Smoreder, 2014 ; Bailliencourt, 2007). En ce qui concerne la modalité de communication en face-à-face, elle est un révélateur majeur de l'intensité de la relation et de la diversité des dispositifs utilisés par les individus pour échanger entre eux. Ainsi, malgré sa perte de centralité dans l'ensemble des modalités relationnelles, l'interaction en face-à-face demeure la modalité de référence pour observer le plus sûrement l'évolution du « *meeting* » (la rencontre physique) au « *mating* » (l'appariement amical) (Verbrugge, 1977 ; Van Dujin, 2003).

Nous retenons que pour appréhender et modéliser le comportement des individus et l'évolution de leurs relations sociales au sein d'une communauté, il apparaît intéressant de prendre en compte le plus possible l'ensemble des actes de communication qu'ils y déploient en même temps et qui narrent une histoire qui s'écrit en entrelaçant plusieurs liens relationnels d'échanges.

Propositions conclusives

Nous concluons cette revue bibliographique avec la proposition suivante : l'observation de l'évolution des relations homophiles vers des relations hétérophiles sur des dimensions exogènes au réseau (les attributs socio-démographies des individus) et endogènes au réseau (les positions structurales) peut ainsi participer à l'appréhension des dynamiques d'évolution d'un réseau social, c'est à dire la formation et la déformation des liens sociaux au sein d'une communauté d'individus homogènes sur certaines caractéristiques et hétérogènes sur d'autres.

Ces relations doivent pouvoir être observables pour opérationnaliser notre étude. Cela est possible sur plusieurs dimensions :

- une dimension déclarative qui révèle les relations affinitaires restituées par l'individu. Cette dimension peut être appréhendée par les réponses à des questions sur les liens dyadiques d'amitié.
- une dimension factuelle qui atteste des contacts dyadiques entre les individus. Elle peut être appréhendée par le recueil de différents actes de communication interpersonnelle.

En outre, nous avons commenté l'importance potentielle des structures d'opportunités et des foyers d'activités dans l'établissement et l'évolution des relations sociales. Ainsi, il apparaît important de pouvoir saisir également le rôle de cette dimension qui médiatise le fait que les individus vivent au sein d'espaces communs et partagent des activités susceptibles de générer des interactions.

Enfin, il nous apparait intéressant de prendre en compte l'entrelacement des observations des liens d'amitiés déclarées, des échanges effectifs et des co-présences des individus.

3. Matériau et méthode

A la section 3.1 de ce chapitre, nous exposons tout d'abord nos hypothèses ainsi que notre proposition pour leur opérationnalisation.

Dans les sections suivantes nous présentons le matériau exploité et la méthode élaborée pour répondre à nos hypothèses. Nous décrivons à la section 3.2 le matériau utilisé, issu d'une importante expérimentation de suivi longitudinal d'une centaine de jeunes professionnels en reprise d'études au MIT. Nous explicitons à la section 3.3 l'originalité de notre étude parmi les travaux déjà réalisés à partir de cette expérimentation. Nous exposons ensuite notre démarche opératoire pour construire les 6 réseaux sociaux uniplexes et multiplexes nécessaires à notre investigation (section 3.4). Enfin à la section 3.5 nous précisons le paramétrage des modèles mobilisés, notamment les variables et les facteurs explicatifs choisis pour répondre à nos hypothèses.

3.1. Formulation des hypothèses et opérationnalisation

A l'appui de la revue bibliographique présentée précédemment et en prémisse de nos hypothèses, nous rappelons plusieurs postulats concernant la dynamique de sélection des liens sociaux que nous étudions :

- dans la poursuite de ses buts, un individu agit comme un acteur « rationnel et raisonnable ».
- l'environnement social de l'individu comprend des foyers d'activité (*foci of activity*), qui constituent des structures d'opportunités au sein desquelles des liens peuvent se créer et évoluer de la rencontre (*meeting*) à l'appariement amical (*mating*).
- le réseau social de l'individu comprend des ressources sociales enchâssées auxquelles il peut potentiellement accéder et qu'il peut mobiliser afin d'atteindre ses buts.
- pour accéder à ces ressources, un investissement relationnel différentiel est nécessaire selon la catégorie des ressources recherchées - expressives ou instrumentales - et leur localisation dans le réseau.
- la dynamique relationnelle de formation et de déformation des liens se déploie progressivement dans le temps. Les effets du passé du réseau au temps t_i sont contenus dans les effets de son état présent observé à un temps t_{i+n}.

Stratégie opératoire

Par des analyses structurales, descriptives et inférentielles, de réseaux sociaux longitudinaux uniplexes et de leurs combinaisons multiplexes, nous souhaitons répondre aux questions suivantes :

1. Dans quelle mesure le phénomène d'homophilie contribue-t-il à expliquer l'évolution de la structure des relations amicales des individus au sein d'une communauté ? Comment les caractéristiques socio-démographiques et structurales qui expliquent l'homophilie co-évoluent-elles dans le temps ?
2. De quelle façon l'expression de l'homophilie se révèle-t-elle, d'une part dans les réseaux uniplexes et, d'autre part, dans les réseaux multiplexes ?
3. Dans quelle mesure les comportements des individus, observés par l'intensité de leurs rencontres et de leurs contacts téléphoniques, traduisent-ils des préférences affinitaires déclarées ?

Pour cela, nous mobiliserons dans un premier temps les outils d'analyse des graphes afin de décrire la dynamique d'évolution des réseaux d'amitiés déclarées, des relations à distance et de proximité. Dans un second temps, nous exploiterons le potentiel des outils de modélisation statistique d'analyse de données longitudinales relationnelles pour des réseaux complets.

Compte-tenu du caractère multi-ethnique de la communauté suivie, nous portons un regard particulier sur l'homophilie relative à l'origine ethnique des individus.

Hypothèse sur l'évolution de l'homophilie socio-démographique (H_homo).

- au temps t_i, au début de la formation de la communauté, l'homophilie des attributs socio-démographiques est le phénomène exogène au réseau social qui explique le plus significativement l'établissement des relations entre les individus. En effet, afin de faciliter leur installation et leur intégration dans la communauté, les ressources les plus recherchées sont les ressources expressives, utiles pour l'intégration dans des lieu et contexte nouveaux. Ces ressources sont détenues par des contacts similaires, du point de vue des attributs socio-démographiques, entre lesquels les premiers liens vont se nouer. Par ce phénomène d'homophilie, la circulation de ces ressources et leur mobilisation est facilitée en matière d'effort d'investissement relationnel.

- au temps t_{i+n}, en cours d'année, l'homophilie relative aux attributs sociaux perd en pouvoir explicatif de formation des liens sociaux. En effet, le but de l'intégration atteint, les individus recherchent progressivement des ressources autres que celles déjà accessibles, ceci afin d'atteindre d'autres buts, plus instrumentaux.

Ces ressources, plus instrumentales qu'expressives, sont détenues par des « Alter » dont les caractéristiques socio-démographiques sont moins similaires que celles détenues par « Ego ».

Un investissement relationnel est nécessaire pour créer des liens avec des individus dont les attributs socio-démographiques sont dissimilaires. Ces liens se nouent en substitution ou en addition de liens déjà établis au temps t_i. Par conséquent, le réseau évolue vers une plus grande hétérogénéité des attributs socio-démographiques. Les effets des attributs qui expriment l'homophilie seront moins significatifs dans l'explication de l'évolution du réseau.

Hypothèses sur l'impact des effets structuraux intra-réseau (H_s1, H_s2, H_s3).

Des phénomènes structuraux endogènes participent à révéler l'investissement relationnel des individus pour créer des liens avec des pairs dissimilaires. Ceci peut être appréhendé par des facteurs structuraux du réseau social, tels que :

- l'activité et la popularité structurale des individus (hypothèse H_s1),
- l'appariement des individus avec des pairs structuralement dissimilaires (H_s2),
- la transitivité relationnelle (H_s3).

Au fil de l'évolution du réseau, ces caractéristiques structurales endogènes, qui peuvent faciliter la mobilisation des ressources sociales recherchées, jouent un rôle proéminent. Par conséquent, les facteurs structuraux endogènes seront significatifs au détriment notamment des facteurs exogènes liés à l'homophilie des attributs dont ils vont réduire la significativité sur l'évolution du réseau d'amitié.

Hypothèses sur les effets structuraux multiplexes (H_s4, H_s5)

L'évolution des relations amicales déclarées correspond en partie seulement à l'évolution des liens sociaux entre les individus. L'observation des échanges sociaux effectifs entre les individus à distance (par téléphone) et en présentiel apporte un éclairage sur l'évolution des relations amicales. Ces échanges répondent à d'autres logiques que celles des déclarations d'amitiés, des considérations plus pratiques et instrumentales. Ces liens se concrétisent au sein de structures d'opportunités qui émergent au fil du temps lors de rencontres et de partage d'activités. Cet entrelacement des relations déclarées et effectives peut être appréhendé par la multiplexité des réseaux qui les traduisent. Ces réseaux s'influencent mutuellement de façon significative dans leur dynamique d'évolution temporelle par des effets structuraux multiplexes, tels que :

- l'impact des choix relationnels des individus au sein d'un réseau sur leurs choix au sein d'un autre réseau (hypothèse H_s4),
- la concordance ou discordance des choix relationnels d'un réseau à l'autre (H_s5).

3.2. Présentation de l'expérimentation *« Friends and Family »*

Introduction

L'expérimentation « *Friends and Family* » menée de 2010 à 2011 par l'Institut Technologique du Massachussets (MIT, Cambridge, Etats-Unis d'Amérique), a produit une base de données relationnelles, unique pour sa richesse, son volume et sa durée, sur les interactions sociales entre des individus (Aharony *et al.*, 2011).

Cette expérimentation a été mise en œuvre afin d'étudier comment des individus forment des choix et prennent des décisions en prenant en compte les relations sociales. A cette fin, une équipe de chercheurs du MIT a réalisé un suivi longitudinal d'une centaine de personnes qui ont quitté leur entreprise pour une reprise d'études au MIT. Ces individus vivent temporairement dans une même résidence étudiante familiale du MIT, nommée *« Westgate Family Housing Community »*. Cette communauté est constituée au total d'environ 400 couples, la moitié avec un enfant et dont au moins un membre du couple suit des études au MIT. Parmi ceux-ci, 130 d'individus représentatifs sont suivis quotidiennement en ce qui concerne notamment leurs interactions sociales. Certaines bases de données concernant cette expérimentation sont accessibles librement, sous réserve d'acceptation d'une demande adressée au MIT, d'autres ne sont pas accessibles librement.

3.2.1. La base de données *« Friends and Family date set »*

Des données sur les relations sociales d'une communauté circonscrite

Les données collectées tout au long de cette expérimentation sont réunies au sein d'une base de données intitulée « *Friends and Family data set* ». Ce riche matériau permet de réaliser une étude des relations de sociabilité entre une centaine d'individus adultes, au sein d'une communauté circonscrite, sur des dimensions géographiques, personnelles et professionnelles :

- un même lieu de vie et d'études : la résidence étudiante familiale du MIT.
- des profils socio-démographiques homogènes sur certains aspects, hétérogènes sur d'autres.
- une même étape dans le parcours professionnel : des individus en reprise d'études supérieures, après avoir exercé une activité professionnelle.

La population de cette communauté est homogène en ce qui concerne le niveau d'études (tous détiennent un diplôme universitaire- *university degree*) et le niveau de vie (données non dsiponibles). La population est hétérogène en ce qui concerne les âges, le nombre d'enfants, l'origine ethnique et la religion. Les participants sont âgés de 22 à 42 ans, avec une moyenne d'âge d'environ 28 ans (données détaillées non disponibles). Huit origines ethniques sont déclarées selon les catégories suivantes : amérindienne, asiatique, blanche, hispanique, israélienne, noire, moyen-orientale, métissée (données détaillées disponibles). Les religions les plus citées sont les suivantes (données détaillées non disponibles) : bouddhiste, catholique, chrétienne, hindoue, juive, mormone. En ce concerne le genre, la population est équilibrée (données détaillées disponibles). Enfin, le nombre d'enfants déclarés varie de 0 à 4 (données détaillées disponibles).

Des données sur l'évolution des relations sociales et amicales dans le temps

Parmi les données disponibles, certaines permettent de réaliser un suivi longitudinal des relations sociales entre les individus. En effet, à l'aide d'un dispositif mis en place par un département de recherche du MIT (*MIT - Media Lab*), environ cent individus représentatifs de cette communauté ont été suivis quotidiennement en ce qui concerne, d'une part, leurs liens d'amitiés, et, d'autre part, leurs contacts interpersonnels distanciels et présentiels.

Les liens d'amitiés sont interrogés à intervalles réguliers au cours de l'année universitaire par des questionnaires auto-administrés. Ces questionnaires sont soumis aux individus suivis afin de recueillir des informations l'évolution de l'intensité de leurs relations amicales les uns envers les autres au sein de cette communauté. Les interactions sociales en co-présence et par téléphone sont suivies à l'aide d'une application, qui utilise le potentiel de la technologie « *Bluetooth* », installée sur les téléphones portables des individus. Un suivi longitudinal est réalisé pendant 9 mois pour les 130 individus de la communauté. La chronologie reconstituée est schématisée dans la suite du chapitre (figure n°2, p. 60)

Plusieurs travaux de recherche ont exploité cette base de données. Dans la partie section suivante, nous présentons une synthèse des problématiques et des principaux résultats issus de la dizaine d'articles à caractère scientifique produits à partir de ce matériau.

3.2.2. Synthèse des études publiées sur l'expérimentation « *Friends and Family* »

Introduction

Une dizaine de travaux, variés en thèmes et méthodes, ont exploité la base de données issue de l'expérimentation « *Friends and Family* » menée par le MIT en 2010-2011 (Aharony *et al.*, 2011). Une partie de ces travaux étudient l'influence des pairs sur les choix des individus avec lesquels ils sont en relation amicale, en ce qui concerne l'achat d'applications pour téléphones mobiles (*mobile apps*) ou la pratique d'une activité physique et sportive. D'autres études s'intéressent à la possibilité de prédire l'humeur, le comportement, des traits psychologiques ou encore différents attributs socio-démographiques des individus de la communauté observée, à partir d'algorithmes d'apprentissage automatique (*machine learning*). Enfin, quelques travaux exploitent des indicateurs structuraux des réseaux constitués.

Première série de travaux : l'influence des pairs sur les comportements d'achat

Une première série de travaux de recherche étudie l'influence des pairs sur les comportements d'achat d'applications pour téléphones mobiles et sur la pratique d'une activité physique.

L'équipe de Pan *et al.* (2011a) s'intéresse à la possibilité de prédire l'installation d'applications de téléphonie mobile à partir des informations récoltées au sein de quatre réseaux sociaux : 1. un réseau constitué à partir du questionnaire sur les liens d'amitiés entre les individus de la communauté ; 2. un réseau constitué par les contacts téléphoniques entre les individus de la communauté; 3. un réseau constitué à partir des relevés des rencontres physiques des individus de la communauté (détectées par la technologie *bluetooth*), et 4. un réseau constitué par l'appartenance commune des individus à un des départements du MIT. Les réseaux 2 et 3 sont construits à partir de relevés cumulés sur une période d'observation de 5 mois. Les réseaux sont ici considérés comme non-dirigés, avec une pondération des liens basée, selon le cas sur : la déclaration d'un lien d'amitié, le nombre d'appels téléphoniques, le nombre de détections de co-présences ou l'appartenance commune des individus à un des départements du MIT.

Les auteurs souhaitent tester deux hypothèses : d'une part, il s'agit d'apprécier si la variance individuelle du comportement des individus en ce qui concerne l'installation d'applications mobiles n'est pas trop importante pour observer un éventuel effet résultant des réseaux de sociabilité.

D'autre part, l'impact de facteurs exogènes, en particulier la « popularité » des applications pour téléphone mobile, sur la qualité de la prédiction des comportements est évalué. A partir d'éléments structuraux et d'informations collectées sur l'installation d'applications pour téléphone mobile par les individus suivis, Pan *et al.* (2011a) développent un modèle algorithmique capable de prédire la probabilité qu'un individu de la communauté installe ou non une application téléphonique particulière.

Plus précisément, parmi les quatre réseaux constitués, les auteurs observent que les données issues du réseau des face-à-face entre les individus donnent les meilleurs résultats pour prédire la diffusion des applications mobiles entre ces individus. En revanche, les données collectées à partir du réseau construit sur la base du questionnaire produisent les moins bonnes prédictions.

Par ailleurs, en combinant les données issues des quatre réseaux sociaux en un seul réseau composite et global, les auteurs parviennent à maximiser la précision des prédictions. Pour cela, ils intègrent d'une part, un facteur exogène - la popularité des applications mobiles - et, d'autre part, un indicateur qui exprime la variance individuelle du comportement des individus pour l'installation d'applications de téléphonie mobile.

Les auteurs mettent en œuvre ici une approche algorithmique uniquement prédictive. Cette modélisation ne permet ni d'étudier les ressorts de la causalité des phénomènes en jeu, ni de décrire les effets de l'influence du réseau de pairs sur l'installation d'applications de téléphonie mobile. Les prédictions sont globalement bonnes, mais le modèle ne permet pas d'expliquer pourquoi les individus adoptent ou non les mêmes choix. En outre, les auteurs soulignent que les mécanismes de diffusion dans le temps des applications mobiles sont très complexes ce qui, selon eux, expliquerait qu'ils n'obtiennent pas de résultats probants en adoptant un modèle qui intègre une variable temporelle.

Les travaux menés par les équipes de Aharony *et al.* (2011) explorent l'influence des pairs sur la pratique d'une activité physique et sportive. Cette recherche a été menée dans le cadre d'un volet spécifique de l'expérimentation globale intitulé « *FunFit experiment* ». Ainsi, les chercheurs mettent en œuvre différentes stratégies expérimentales afin de mieux comprendre l'impact de l'influence sociale exercée par les amis (*peer buddies*) sur la pratique d'une activité physique d'intensité croissante, en incluant une gratification numéraire proportionnelle à la performance de l'individu ou de ses pairs.

A partir des proximités amicales déclarées via un questionnaire adressé à une centaine d'individus de la communauté, les auteurs constituent un réseau de proximité sociale quasi-complet (*network of self perceived closenness*). L'expérimentation consiste alors à constituer des groupes formés de

triades de connaissances amicales plus ou moins proches : 1. un premier groupe, dit « de contrôle », au sein duquel les individus sont récompensés financièrement en fonction de leurs propres progrès d'intensification hebdomadaire de leur pratique sportive ; 2. un deuxième groupe, dénommé « *peer-see* », au sein duquel chaque individu a connaissance des progrès des deux autres membres de la triade. Dans cette configuration, chaque individu connait la performance des deux autres individus de la triade et est récompensé financièrement uniquement en fonction de ses propres progrès ; 3. un troisième groupe, intitulé *« peer-reward »*, au sein duquel chaque individu a connaissance des progrès des deux autres membres de la triade.

Dans ce groupe, chacun est récompensé financièrement uniquement en fonction des progrès des deux autres membres de la triade. L'expérimentation menée au sein du troisième groupe s'inscrit dans le cadre théorique des mécanismes sociaux basés sur les interactions et la pression sociale entre pairs (*peer-to-peer interactions and peer pressure*), développé entre autres par Mani *et al.* (2010, cités par Aharony *et al.*, 2011).

Les résultats de cette étude montrent un effet statistiquement significatif de l'influence des pairs sur le comportement des individus dans l'adoption d'une activité physique régulière et d'intensité croissante. Plus précisément, l'étude montre une influence différentielle des pairs selon leur niveau de proximité amicale. Ceci suggère aux auteurs un phénomène complexe de « contagion » au sein des triades formées par les individus, tel que décrit par Centola et Macy (2007, cités par Aharony *et al.*, 2011).

Dans cette étude, la constitution du réseau d'amitié déclaratif sert uniquement à construire les 3 groupes de triades. Les éléments proprement « structuraux » du réseau social ne sont pas exploités par les auteurs.

Sur le même thème, Almaatouq *et al.* (2016a, 2016b) étudient les effets de la réciprocité et de la direction des liens d'amitiés sur l'influence entre pairs (*social influence and peer-pressure*). Les auteurs formulent l'hypothèse que les indices structuraux de réciprocité et de direction des liens du réseau d'amitiés constituent des facteurs essentiels dans la compréhension des mécanismes d'influence entre pairs, au-delà de l'effet dû à l'existence d'un lien déclaré. En réinvestissant l'expérimentation menée par Aharony *et al.* (2011), les auteurs souhaitent explorer les changements de niveau d'investissement d'« Ego » dans une pratique sportive lorsqu'il est influencé par des « Alters » perçus par « Ego » comme des connaissances plus ou moins proches.

Les auteurs mettent en œuvre une modélisation par régression avec, d'une part, en variable dépendante les changements de comportement d'« Ego » dans sa pratique d'une activité physique et, d'autre part, trois variables indépendantes qui permettent d'appréhender la direction des liens entre « Ego » et ses « Alters » : 1. l'existence d'un lien réciproque, 2. l'existence d'un lien dirigé d'« Ego » vers « Alter », 3. l'existence un lien dirigé d'« Alter » vers « Ego ».

Les résultats obtenus montrent que des changements statistiquement significatifs dans les comportements adviennent quand les individus partagent des liens réciproques ou lorsqu'il existe un lien d'amitié dirigé et non réciproque de l'individu qui exerce une pression sociale (Alter) vers l'individu cible (Ego).

Cependant, aucun changement statistiquement significatif n'est perceptible par le modèle de régression utilisé lorsque le lien d'amitié est dirigé de l'individu cible vers l'individu exerçant une pression sociale. Par ailleurs, les variables de contrôle du modèle sont : 1. la force des liens, dont l'effet modérateur dans les phénomènes d'influence sociale a été souligné par Aral et Walker (2014, cités par Almaatouq *et al.*, 2016a et 2016b) ; 2. le genre et l'âge d'« Ego » ; 3. les similitudes de genres et d'origines ethniques entre « Ego » et « Alter ». Ces résultats apparaissent différer sensiblement avec ceux obtenus par d'autres études sur les effets de la « contagion sociale », tels que celles réalisées par Christakis et Fowler sur la diffusion des pratiques alimentaires (2007) et du sentiment de bonheur (2009, cités par Almaatouq *et al.*, 2016a et 2016b).

Les travaux d'Almaatouq *et al.* (2016a, 2016b) mettent également en évidence une incapacité importante des individus à percevoir de façon correcte la réciprocité des liens d'amitiés. Les résultats obtenus soulignent l'enjeu de l'image de soi dans l'asymétrie qui apparaît au niveau des liens déclarés réciproques. Ceci semble avoir des conséquences importantes sur la capacité des individus à influencer un pair et à modifier son comportement. Les auteurs montrent que l'effet de la direction des liens est plus important que l'effet de l'intensité déclarée des liens, qu'ils désignent comme constitutifs d'un « *capital social implicite* » (Almaatouq *et al.*, 2016b, p.34). Enfin, les auteurs révèlent que deux indicateurs structuraux du réseau d'amitiés entre les individus - l'encastrement (*social embeddedness*, Granovetter, 1985) et la centralité (*social centrality)* - peuvent efficacement prédire la direction et la réciprocité des liens entre deux individus. Ainsi, dans une certaine mesure, ces indicateurs peuvent servir de proxy en compensant l'absence d'information sur la direction des liens, dont les effets sur les phénomènes d'influence entre pairs ne peuvent pas être ignorés.

Deuxième série de travaux : prédictions des attributs socio-démographiques des individus

Une deuxième série d'études tente de prédire des descripteurs tels que la personnalité des individus, leur état de santé, leur origine ethnique ou encore leur nationalité, par des indices structuraux des réseaux sociaux formés par ces individus (Staiano *et al.*, 2012) ou par des indicateurs construits à partir des contacts entre les individus (Altshuler *et al.*, 2011 ; Moturu *et al.* 2011 ; Montjoye *et al.*, 2013).

Ainsi, Staiano *et al.* (2012) interrogent la relation entre la structure de trois réseaux sociaux formés à partir d'observations agrégées sur une période de trois mois et des traits de personnalité d'un groupe de cinquante-cinq individus. D'une part, un premier réseau est construit à partir des rencontres de proximité, un deuxième est basé sur les contacts téléphoniques et enfin un dernier est établi à partir du questionnaire sur les liens d'amitiés entre les individus de la communauté.

D'autre part, la personnalité des individus de ce gourpe est appréhendée par cinq traits psychologiques caractéristiques issus du « *big five personality test* » (Goldberg, 1971, cité par John et Srivastava, 1999). Les auteurs privilégient dans cette étude l'exploitation des caractéristiques structurales des réseaux sociaux et comparent les résultats avec ceux obtenus par l'approche, plus classique, qui consiste à exploiter les attributs des acteurs (*actor-based features*), tels que le nombre d'appels et la durée moyenne des appels pour chaque individu. Plusieurs indicateurs structuraux sont étudiés, notamment la centralité, le recensement des triades (*triad census*, David et Leinhardt, 1972, cité par Staiano *et al.*, 2012), la transitivité et l'indice d'efficacité de la propagation de l'information au sein de formes de type « petits mondes » (*small world and efficiency measure*, Latora et Marchiori, 2001, cité par Staiano *et al.*, 2012), des indices structuraux centrés sur les « *ego network* » (transitivité et efficience locale, triades). Les auteurs obtiennent des prédictions des traits psychologiques des individus d'une meilleure qualité à partir des indices structuraux qu'à partir des attributs individuels. En outre, les prédictions ne sont pas de même qualité selon le type de réseau. Ainsi, le réseau basé sur le questionnaire apparaît le moins performant pour prédire les traits de personnalité des individus.

Dans un même registre, Montjoye *et al.* (2013) entreprennent de prédire les traits de personnalité d'une cinquantaine d'individus de la communauté observée, cette fois à partir des appels téléphoniques et des sms cumulés sur une année.

Pour cela, ils mettent en œuvre des algorithmes du type « séparateurs à vaste marge » (*support vector machine*, Boser *et al.*, 1992), fondés sur un ensemble de techniques algorithmiques « d'apprentissage supervisé ». Avec un taux de prédiction correct, variant de 51 % à 61%, les algorithmes utilisés sur les données permettent de prédire différents traits de personnalité des individus, notamment l'extraversion, l'amabilité, l'esprit de responsabilité ou encore l'ouverture d'esprit.

Pour cela, les auteurs exploitent cinq indicateurs basés sur les appels téléphoniques et les sms échangés : 1. le nombre d'appels/sms ; 2. le comportement actif de l'individu (le nombre d'appels initiés et le temps de réponse à un sms reçu) ; 3. la régularité (la routine temporelle des appels et l'intervalle temporel des appels/sms) ; 4. la diversité (le ratio du nombre d'appels et du nombre de détections de rencontres en face-à-face). Certains de ces indicateurs enrichissent l'appréhension de l'intensité des liens entre les individus par la prise en compte de leur évolution dans le temps. Toutefois, cette étude ne fait pas état d'une exploitation des caractéristiques structurales des réseaux et, comme pour les travaux de Staiano *et al.* (2012), l'approche est prédictive et non explicative.

Altshuler *et al.* (2011) développent des modèles prédictifs de différents attributs des individus de la communauté : l'origine ethnique, la nationalité, le statut d'étudiant et la présence d'un conjoint. Cette modélisation est réalisée à partir des interactions observées au sein des réseaux sociaux formés à partir des échanges de sms et de la détection des rencontres en face-à-face, sur une période d'un mois. Les auteurs utilisent à cette fin des algorithmes d'apprentissage automatique (*machine learning*) et étudient l'évolution du processus d'apprentissage à mesure que l'information disponible augmente. Les auteurs obtiennent des résultats statistiquement significatifs avec une qualité de prédiction supérieure à 60% pour les quatre attributs « classificateurs » sélectionnés : l'origine ethnique (classement obtenu par une détection de communauté basée sur l'algorithme de Louvain appliqué à l'échange de sms), la nationalité (*US-natives* ou non), le statut d'étudiant et la présence d'un conjoint *(life-partners*).

Moturu *et al.* (2011) utilisent la richesse de la base de données pour étudier les liens entre la sociabilité des individus de la communauté et leur état de santé. Dans cette étude, la sociabilité des individus est appréhendée par le nombre de contacts en face-à-face cumulés sur un mois. L'état de santé est mesuré par l'humeur des individus et leur durée de sommeil, telles que déclarées dans un questionnaire administré quotidiennement.

Les auteurs confirment l'existence, bien connue par ailleurs, de relations statistiquement significatives entre la durée de sommeil et l'humeur des individus. En outre, les auteurs observent, d'une part, que les individus ayant la plus faible « sociabilité » mentionnent être de mauvaise humeur (*poor mood*) plus fréquemment et de façon statistiquement significative. D'autre part, il est observé une association significative entre la durée quotidienne de sommeil des individus et leur sociabilité. Dans cette étude, les éléments structuraux des réseaux sociaux ne sont pas exploités.

Troisième série de travaux : diversité des relations sociales et niveau de revenus

Les dernières études explorent les liens entre la diversité des relations sociales des individus et leur niveau de revenus, à partir de données non structurales (Pan *et al.*, 2011b). Enfin, les travaux les plus récents rendent compte de l'élaboration de modèles capables de prédire le niveau de confiance des individus les uns envers les autres, en utilisant des éléments structuraux et des algorithmes d'apprentissage automatique (*Shmueli et al., 2014 ; Nishikata et al., 2018).*

Les relations entre la diversité des interactions sociales et les revenus des individus de la communauté sont étudiées par Pan *et al.* (2011b). Les auteurs exploitent le nombre total d'appels téléphoniques et de rencontres en face-à-face entre les individus pendant plusieurs mois, afin de rendre compte de la distribution du temps qu'un individu consacre aux autres. Leurs résultats montrent que les individus qui disposent d'un capital économique élevé manifestent des comportements sociaux, en termes de diversité des interactions sociales, significativement différents de ceux qui disposent d'un moindre capital économique. Ainsi, par exemple, la diversité des contacts téléphoniques et la diversité des contacts de proximité sont significativement corrélées avec le niveau courant de revenus, ce qui n'est pas le cas du nombre total d'appels ou de la durée des rencontres physiques. Ainsi, les individus économiquement aisés n'ont pas forcément plus de contacts, mais les répartissent plus uniformément parmi leurs relations.

A partir de leurs résultats, Pan *et al.* (2011b) interrogent certains aspects du cadre théorique qui soutient que la diversité des interactions sociales à un impact positif sur l'élévation du capital économique (Burt, 1995, Seidel *et al.*, 2000, Bruggeman, 2010, cités par Pan *et al.*, 2011b).

Selon cette perspective en effet, les individus qui seraient soudainement privés de leurs revenus élevés, par exemple pendant une période de reprise d'études (ce qui est la situation des individus de la communauté observée dans notre propre étude), continueraient à entretenir une large diversité d'interactions sociales, ce comportement ayant contribué à l'augmentation de leur capital économique. Or, l'étude de Pan *et al.* (2011b) montre que la diversité des interactions sociales courante n'est pas corrélée significativement avec le capital économique passé : au contraire, il apparaît seulement une corrélation significative entre les revenus courants et la diversité courante des interactions sociales. Ainsi, selon ces résultats, les individus ne profiteraient pas systématiquement de leurs expériences passées, en termes de bénéfice éprouvé de la diversification de leurs interactions sociales, dans une situation nouvelle où leur capital économique est temporairement plus faible.

Toutefois, l'étude de ces auteurs ne prend pas en compte la variable temporelle. Or, on peut raisonnablement faire l'hypothèse qu'une approche longitudinale sur plusieurs mois apporterait un éclairage intéressant sur cette problématique, par exemple en étudiant la corrélation entre le différentiel de niveau de revenus passés-courants et la dynamique de l'évolution de la diversité des interactions sociales des individus dans le temps. Hélas, le revenu des individus suivis n'est pas fourni dans la base de données en accès libre de l'expérimentation « *Friend and Family* ».

L'équipe de Shmueli et al. (2014) s'intéresse à la confiance, en tant qu'élément constitutif des interactions sociales et économiques, entre une cinquantaine d'individus. Les auteurs forment l'hypothèse que la confiance interpersonnelle facilite l'accès à des ressources sociales et constitue ainsi un capital social (Lin 1982, 1995 ; Coleman, 1988 ; Burt, 2000). En outre, la confiance interpersonnelle est supposée avoir un effet sur la persuasion entre pairs. La définition de la confiance étant différente selon les disciplines des sciences humaines et sociales, les auteurs opérationnalisent le concept de confiance à l'aide de trois questions posées aux individus sur trois domaines qui impliquent une forte relation de confiance : 1. les finances (une demande de prêt d'argent) ; 2. la famille (une garde d'enfant) ; 3. la santé (une aide lors d'une maladie). A cette fin, les auteurs utilisent des éléments structuraux des réseaux sociaux construits à partir des appels téléphoniques, des sms et des rencontres en face-à-face. A l'aide de six indicateurs structuraux, dont plusieurs indices de centralité, et d'un algorithme d'apprentissage automatique, les auteurs obtiennent des prédictions correctes (à plus 90%) des réponses données par les individus.

Dans cette étude, la passation du questionnaire qui opérationnalise la confiance est réalisé une seule fois, au début de l'expérimentation « *Friends and Family* ». Ce fait ne permet pas une étude dynamique du phénomène, ni une étude de l'évolution du sentiment de confiance entre les individus au cours du temps.

Enfin, Nishikata et al. (2018) proposent une mesure du capital social à partir des données relationnelles de la base de données « Friends and Family data set ». Dans cette étude, les auteurs construisent ce qu'ils nomment « unité de capital social » à partir des éléments suivants : chaque fois que deux individus ont des interactions réciproques, une unité de capital social est comptée. Les interactions considérées ici sont celles qui impliquent une confiance interpersonnelle. Par exemple, il peut s'agir d'une requête adressée à un pair pour obtenir un service ou une aide financière. Les interactions sont observées à l'aide du réseau social formé par les appels téléphoniques, dont Shumeli *et al.* (2014) ont montré l'intérêt pour mesurer efficacement la confiance entre des individus. Les auteurs construisent ainsi un réseau social qui reflète la « distribution du capital social » au sein de la communauté des individus.

A l'aide d'un algorithme d'apprentissage automatique ils obtiennent des prédictions correctes (à plus de 90%) des réponses des individus à des questions qui impliquent une forte relation de confiance sur trois domaines : les finances, la famille et la santé. Ces auteurs utilisent uniquement la variable construite « unité de capital social » détenue par les individus, sans exploiter les indicateurs structuraux du réseau. En outre, la dimension temporelle n'est pas prise en compte : les « unités de capital social » sont calculées uniquement sur la période totale de l'expérimentation et l'évolution de cet indicateur en fonction de l'évolution temporelle des interactions observées au sein du réseau n'est pas exploitée.

3.3 Positionnement et originalité de la présente étude

La dizaine d'études publiées présentées ci-avant, qui exploitent les données de la base constituée par l'expérimentation « *Friends and Family* », sont très variées en thèmes et méthodes de traitement des données. Nous observons toutefois que parmi ces travaux peu exploitent les éléments structuraux des réseaux sociaux qui peuvent être formés par les observations répétées dans le temps des relations entre les individus de la communauté suivie. Le cas échant, quelques éléments structuraux sont utilisés à des fins prédictives pour entrainer des algorithmes du type *« machine learning »*. Ces derniers sont performants pour la prédiction mais ils sont muets en ce qui relève de l'explication des phénomènes sociaux en jeu.

Ainsi, à ce jour et à notre connaissance, aucune recherche publiée ne fait état d'une exploitation des données de l'expérimentation « *Friends and Family* » dans l'objectif d'une analyse longitudinale des réseaux sociaux, uniplexes et multiplexes. Par conséquent, notre étude apparaît originale dans l'exploitation des données issues de cette expérimentation. A partir de l'évaluation du potentiel des données disponibles, cette expérimentation nous apparaît un matériau intéressant pour étudier l'évolution dans le temps des liens de sociabilité et du phénomène d'homophilie au sein d'un réseau complet d'individus.

Dans la section suivante, nous présentons le processus d'apurement, de documentation et d'agencement des données relationnelles que nous utilisons pour élaborer les réseaux uniplexes et multiplexes formés par les liens d'amitiés déclarés (données des questionnaires auto-administrés), les échanges distanciels (données sur les contacts téléphoniques) et présentiels.

3.3.1 Apurement et agencement des bases de données relationnelles

Préalablement à la construction des réseaux sociaux utilisés pour notre étude, un travail d'apurement a été nécessaire afin de préparer les jeux de données relationnelles utiles pour notre étude. Ce travail, nommé couramment sous le terme « *data wrangling* »[4], peut-être décrit comme le processus qui permet, à partir de données brutes ou nettoyées, de les structurer, les enrichir et les documenter de manière la plus exhaustive possible.

[4] « Le travail de '*data wrangling'* est aussi appelé 'préparation de données en self-service'. Le '*data wrangler*' peut être vu comme le cowboy (nommé *'wrangler'* au nord des Etats-Unis d'Amérique) de la donnée, qui essaie de rassembler ses données éparpillées de la même manière qu'un cowboy rassemble ou tri son bétail. Le terme '*wrangling*', utilisé dans le langage courant aux Etats-Unis, suggère que l'activité est parfois laborieuse, souvent chronophage, mais que celle-ci doit nécessairement être réalisée afin d'aboutir à un travail bien fait ». Dans notre étude, le cheval s'appelle « R ». Cette définition est librement adaptée de celle proposée par le site internet : www.trifacta.com/blog/quest-ce-que-le-data-wrangling (consulté en mars 2020).

Ce travail fait partie d'un processus global de préparation, d'identification et d'ordonnancement des données qui s'est développé et structuré au fur et à mesure de la croissance des grandes bases de données et par la nécessité de mise en œuvre de techniques fiables de préparation des données préalablement à leur exploitation.

Les données en libre accès fournies par le laboratoire du MIT à l'origine de l'expérimentation (*MIT Human Dynamics Lab)*[5], se présentent sous la forme de fichiers de formats, de tailles et de structures variables. Le volume et l'ordonnancement des données a nécessité un travail préparatoire assez conséquent. Ainsi, par exemple, le fichier le plus volumineux que nous avons traité contient plus de 6 millions d'observations * 6 variables (fichier de collecte des contacts quotidiens en face-à-face). Il ne peut pas être ouvert en entier sous Excel (*Microsoft Office Professionnel Plus 2016*). Les fichiers comportant ces données nous ont conduit à développer des programmes spécifiques avec le langage de programmation statistique et de sciences des données « R » (Ihaka et Gentleman, 1996), afin de pouvoir les manipuler aisément en format matriciel et les exploiter.

Afin de faciliter l'agencement des données, nous avons élaboré une frise chronologique qui présente une restitution synthétique et exhaustive de l'organisation du recueil des données de l'expérimentation « *Friends and Family* » (figure n°2, p.60). Aucun document de ce type n'étant cité ni mis à disposition par le laboratoire MIT Media Lab, il s'agit par conséquent d'une production originale qui synthétise l'essentiel de l'expérimentation et situe notre propre étude.

Nous avons élaboré cette frise à partir de plusieurs matériaux : 1- les informations en libre accès fournies par le laboratoire du MIT à l'origine de l'expérimentation ; 2- l'exploitation des articles scientifiques synthétisés dans le chapitre précédent ; 3- des prises de contacts directs avec quatre chercheurs[6] qui ont participé à la mise en œuvre de l'expérimentation, afin recueillir des informations complémentaires et préciser certains points.

[5] Les données de l'expérimentation « *Friends and Family* » (Aharony *et al.*, 2011) sont en libre accès sur le site internet du MIT : http://realitycommons.media.mit.edu/friendsdataset.html.

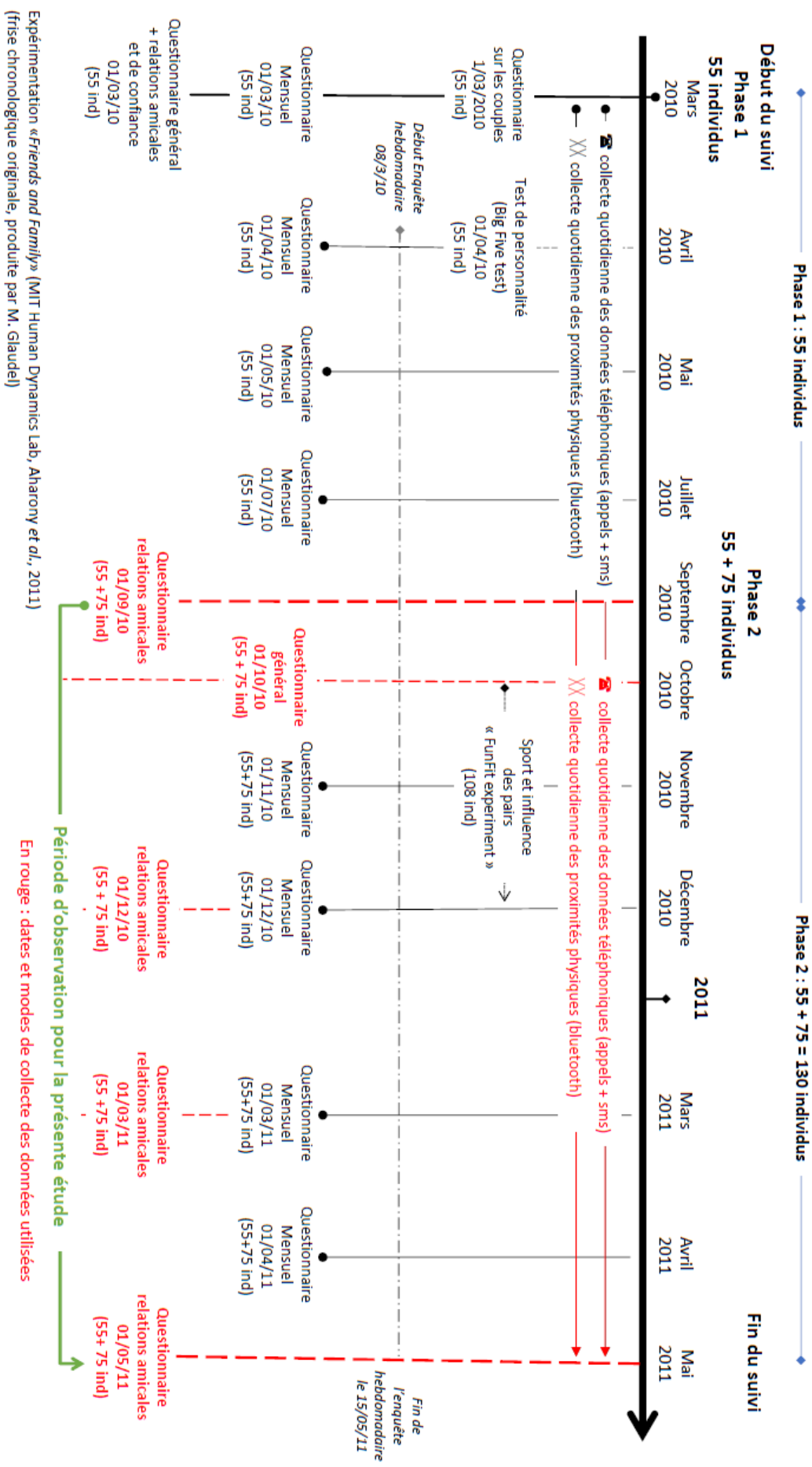
[6] Je remercie les chercheurs contactés pour les compléments d'information : messieurs Abdullah M. Almaatouq (Computational Social Science – MIT), Bruno Lepri (visiting scientist - MIT Media Lab), Erez Shmueli (senior lecturer, head of the Big Data Lab - Department of Industrial Engineering, Tel-Aviv University ; research affiliate - MIT Media Lab), Vivek K. Singh (associate professor, School of Communication and Information, Rutgers University, visiting professor, research affiliate - MIT Media Lab).

Finalement, après d'une part, un travail de construction d'une documentation étendue rendant compte de la chronologie de l'expérimentation et, d'autre part, une exploration approfondie des données, nous avons apuré et agencé un jeu de données qui nous permet de construire des réseaux sociaux complets, avec peu d'attrition. Celui-ci comprend le suivi d'une centaine d'individus pendant 9 mois à partir duquel nous pouvons construire 3 réseaux sociaux longitudinaux uniplexes :

1. **un réseau « R_Friend », construit à partir des liens d'amitiés déclarées entre les individus suivis via des questionnaires auto-administrés et répétés dans le temps.**
2. **un réseau « R_Call », élaboré à partir des relevés des contacts téléphoniques entre les individus suivis.**
3. **un réseau « R_Prox », bâti à partir de données issues de la détection des co-présences physiques des individus suivis.**

La combinaison des ces réseaux uniplexes, nous permet de construire des réseaux multiplexes qui traduisent l'entrelacement des relations sociales d'amitiés déclarées, avec les relations sociales observées par échanges distanciels (contacts téléphoniques) et présentiels (co-présences physiques). Trois réseaux multiplexes rendent compte de cet entrelacement :

1. **un réseau d'ordre 2 : « R_Friend » ∩ « R_Call**
2. **un réseau d'ordre 2 : « R_Friend » ∩ « R_Prox »**
3. **un réseau d'ordre 3 : « R_Friend » ∩ « R_Call »∩ « R_Prox »**



Expérimentation «*Friends and Family*» (MIT Human Dynamics Lab, Aharony *et al.*, 2011)
(frise chronologique originale, produite par M. Glaudel)

Figure n°2 : chronologie du recueil des données de l'expérimentation « Friends and Family » (Aharony et al., 2011). (Production originale).

3.4. Méthodologie pour l'étude descriptive et longitudinale des réseaux

Dans les sections suivantes, nous détaillons la méthode suivie pour construire les 3 réseaux uniplexes et multiplexes dont l'étude statique et dynamique est susceptible de fournir des réponses à nos hypothèses (h_homo, h_s1 à h_s5) et plus largement à éclairer notre problématique (section 3.5). Nous préciserons ensuite (section 3.6) la méthodologie suivie pour réaliser les simulations de modélisation longitudinales des réseaux uniplexes, d'une part et, d'autre part, des réseaux multiplexes.

L'élaboration des réseaux longitudinaux

Nos bases de données relationnelles sont construites à partir du suivi longitudinal d'un groupe de 130 individus professionnels en reprise d'études au sein d'une grande université américaine (*Massachusetts Institute of Technology*, MIT). Ces individus sont suivis tous les jours de septembre 2010 à mai 2011, dans leurs contacts les uns avec autres à l'aide d'une application installée sur leur téléphone portable.

Les 130 individus suivis forment une communauté : ils partagent un même résidence familiale étudiante, ils sont plutôt jeunes, en couple, environ la moitié des couples a un enfant. Un membre du couple au moins fréquente le MIT et plusieurs d'entre eux ont quitté leur entreprise pour une reprise d'études au MIT. Pour la plupart, ils bénéficient d'une aide financière pour se former et ils exerçaient avant cette reprise d'études « des carrières en réussite avec en moyenne des salaires de haut niveau » (Pan *et al.*, 2011 p. 3).

Les données collectées tout au long du suivi longitudinal de cette communauté permettent d'établir 3 réseaux sociaux et d'observer leurs évolutions durant 9 mois, de septembre 2010 à mai 2011, à partir :

1. des réponses aux 4 questionnaires sur les relations amicales, reconduit par intervalle de 3 mois.
2. du nombre d'appels téléphoniques échangés entre les individus suivis chaque jour.
3. des proximités physiques détectées entre les individus suivis chaque jour.

Le tableau n°1.a synthétise les calendriers des collectes des données des 3 bases relationnelles qui sont utilisées pour former les 3 réseaux sociaux : « R_Friend », « R_Call » et « R_Prox ».

Tableau n° 1.a : chronologie de la collecte des données de l'expérimentation « Friends and Family »

Vague n°1		
Septembre 2010	Octobre 2010	Novembre 2010
Questionnaire amitié n°1 Réseau R_Friend		
Relevé des contacts téléphoniques Réseau « R_Call »		
Relevé des proximités physiques Réseau « R_Prox »		

Vague n°2		
Décembre 2010	Janvier 2011	Février 2011
Questionnaire amitié n°2 Réseau R_Friend		
Relevé des contacts téléphoniques Réseau « R_Call »		
Relevé des proximités physiques Réseau « R_Prox »		

Vague n°3		
Mars 2011	Avril 2011	Mai 2011
Questionnaire amitié n°3 Réseau R_Friend		Questionnaire amitié n°4 Réseau R_Friend
Relevé des contacts téléphoniques Réseau « R_Call »		
Relevé des proximités physiques Réseau « R_Prox »		

3.4.1. Le réseau uniplexe des relations amicales « R_Friend »

Il s'agit ici d'une mise en réseau des données issues de 4 questionnaires auto-administrés sur les relations amicales déclarées entre les individus de la communauté suivie. La période de passation des questionnaires se déroule sur 4 vagues : au mois de septembre 2010 (vague n°1), au mois de décembre 2010 (vague n°2), en mars 2011 (vague n°3) et au mois de mai 2011 (vague n°4). Chaque individu est interrogé au même moment et est invité à noter sa proximité amicale (*friendship*) avec les autres individus de la communauté, sur une échelle d'intensité variant de 0 (aucun lien d'amitié) à 7 (forte proximité amicale). Nous considérons cette information comme un indicateur de la force du lien déclaré d'amitié entre les individus.

Les questions posées sont les suivantes :

« ***0****: I don't know this person ;* ***1****: I know of this person ;* ***2****: This person is an acquaintance ;*

3*: This person is a friend (low ranking) ;* ***4****: This person is a friend (medium ranking) ;*

5*: This person is a friend (high ranking) ;* ***6****: This person is a close friend ;*

7*: This person is family or as close to me as a family member* ».

Notre étude portant sur un réseau de liens d'amitiés, nous ne prenons en compte que les intensités égales ou supérieure à 1, ce qui correspond à la présence d'un lien déclaré entre 2 individus. Nous considérons qu'à partir de l'intensité 3 - *« This person is a friend (low ranking) »* - un lien d'amitié substantif est effectivement déclaré.

A l'examen des données, nous observons des données manquantes concernant les individus et le recueil de données issues des questionnaires. D'une part, certains individus ne sont pas notateurs à toutes les vagues de passation du questionnaire ou ne sont pas toujours notés à chaque vague. D'autre part, certains individus notent des pairs mais ne sont parfois pas notés par ces derniers. Le schéma n°1 présente la variation des effectifs des individus qui ont noté leurs pairs, des vague n°1 à n°4, ainsi que, pour chaque vague considérée, les effectifs des individus qui sont notés par leurs pairs. Selon les vagues, on observe des effectifs variables d'individus répondants (notateurs) et désignés (notés).

Au sein des individus notateurs, selon les vagues de passation du questionnaire, on observe une attrition maximum de 23%, avec des effectifs qui peuvent varier d'un maximum de 117 (vague n°1) à un minimum de 91 (vague n°3). Au sein des individus notés, l'attrition maximale est de 2% seulement, avec un maximum de 126 individus notés et un minimum de 124.

Dans le cadre d'une modélisation longitudinale de réseaux sociaux, telle que nous l'envisageons à l'aide de l'algorithme « Siena », Snijders (2001) et Ripley *et al.* (2020) indiquent un seuil de tolérance de 10 % de données manquantes pour espérer une convergence stable de l'algorithme. Le premier enjeu est par conséquent de pouvoir disposer d'un effectif d'individus, de composition stable dans le temps, c'est à dire avec un minimum d'individus partants (*leavers*) et de nouveaux entrants (*joiners*) de la première à la dernière vague de passation du questionnaire. Le second enjeu est de réduire au minimum le volume de données manquantes dans les réponses aux questionnaires.

Afin de répondre à ces enjeux, nous choisissons tout d'abord d'exclure toutes les données manquantes ainsi que les individus « volatiles » (*leavers* ou *joiners*). Les données manquantes sont ici identifiées comme une absence de réciprocité dans la notation : un individu « A » indique un lien d'amitié vers un individu « B », mais l'information est manquante quant à l'intensité[7] du lien réciproque « B » vers « A ».

A partir de ce choix, nous avons identifié un groupe d'individus permanents (c'est à dire sans partants ni nouveaux entrants) pour les 4 vagues de passation du questionnaire et sans données manquantes. Ce groupe est composé de 70 « individus notateurs » permanents qui notent 123 individus permanents, pour les 4 vagues d'administration du questionnaire sur les liens d'amitiés déclarés.

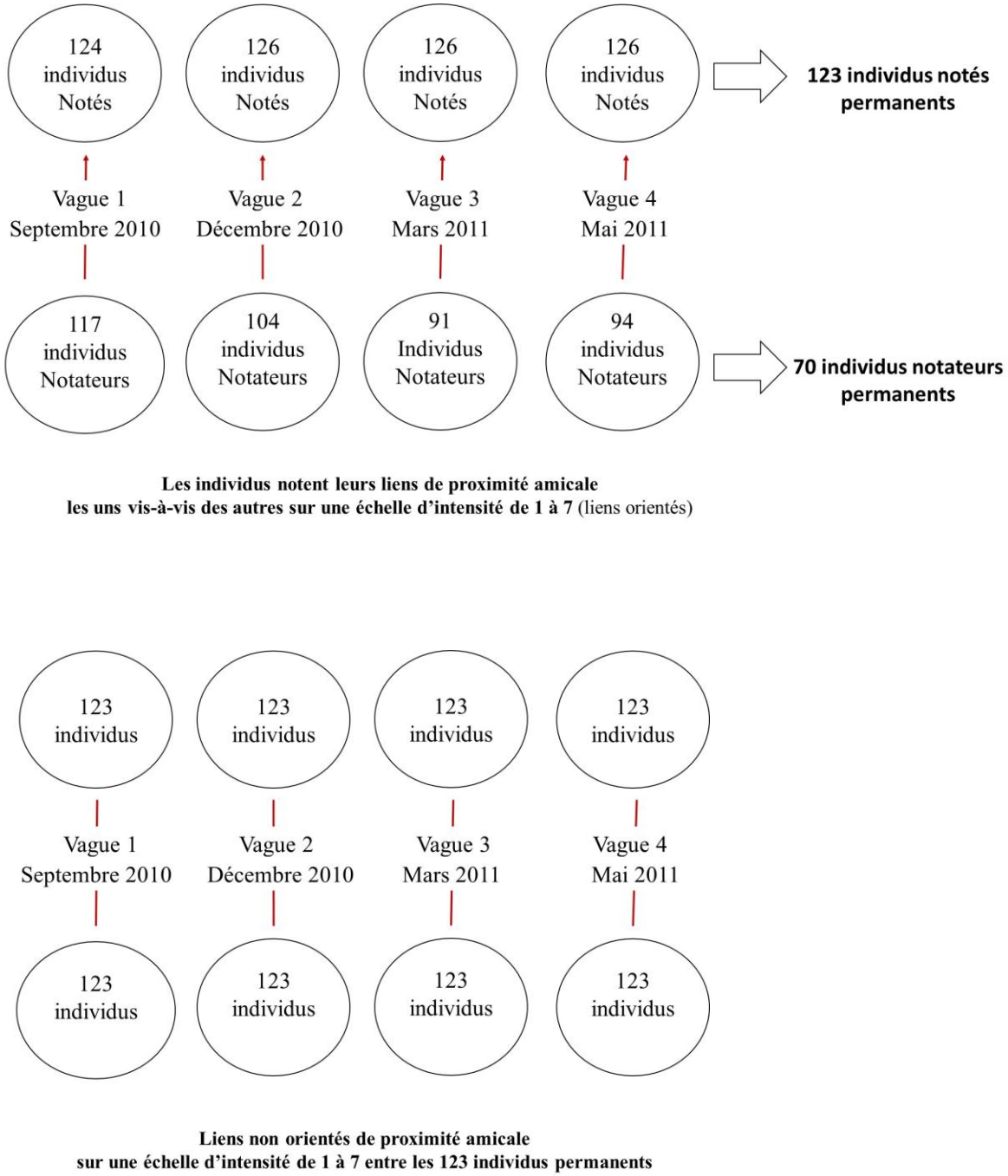
En outre, nous considérons que les liens d'amitiés sont non orientés. Ainsi, si un individu « A » indique un lien d'amitié d'intensité « i » avec un individu « B » et que le lien réciproque est manquant dans la base de données, nous considérons alors que le lien d'amitié déclaré par l'individu « B » vers l'individu « A » est d'intensité « i ». Le lien d'amitié entre l'individu « B » et l'individu « A » est ainsi considéré comme non orienté et d'intensité « i ». De même, si un individu « A » indique un lien d'amitié d'intensité « i » vers un individu « B » et que le lien réciproque est d'une intensité « j », nous considérons un lien d'amitié non orienté entre l'individu « A » et l'individu « B ».

L'intensité retenue pour les liens non orientés est l'intensité la plus élevée. Ce choix est préféré à l'option de la moyenne qui pourrait conduire à atténuer les écarts d'intensité des relations au sein du réseau.

[7] L'intensité des liens varie de la note 0 à 7. Par conséquent, l'absence de note est considérée comme une donnée manquante.

Finalement, nous disposons ainsi de liens non orientés d'amitié déclarée entre 123 individus permanents sur les 4 vagues de passation du questionnaire, à partir desquels nous construisons le réseau « R_Friend » (figure n°3).

Figure n°3 : représentation de la construction d'une forme non orientée du réseau « R_Friend »



3.4.2. Le réseau uniplexe des contacts téléphoniques « R_Call »

A partir des relevés des contacts téléphoniques entre les individus suivis, nous construisons un réseau social de contacts téléphoniques nommé « R_Call ». Pour les individus suivis par l'étude, un relevé quotidien de l'identité de l'appelant et du destinataire, le jour et l'heure de l'envoi sont enregistrés, mais pas le contenu de l'appel. La collecte des données pour la période observée s'étale sur 3 vagues de 3 trimestres, de septembre 2010 à mai 2011. Elle s'inscrit dans les 3 intervalles des 4 enquêtes auto-administrées réalisées sur les relations amicales qui constituent le réseau « R_Friend ». Le nombre de contacts téléphoniques entre les individus est relevé tous les jours, de septembre 2010 à mai 2011. Une première période, nommée « vague 1 » dans le tableau n°1.b, correspond au nombre de contacts téléphoniques entre les individus de septembre 2010 à novembre 2010. Une deuxième période, nommée vague n°2, correspond au nombre de contacts téléphoniques entre les individus entre décembre 2010 et février 2011. Une troisième période, nommée vague n°3, correspond au nombre de contacts téléphoniques entre les individus de mars à mai 2011.

Construction de l'indicateur de lien

Nous faisons l'hypothèse que les contacts téléphoniques sont essentiellement des contacts entre des personnes qui sont dans un rapport de proximité amicale. Nous choisissons d'observer, au sein de chacune des 3 périodes, le nombre d'appels téléphoniques quotidiens et nous synthétisons l'information en construisant un indicateur de lien entre les individus, avec les choix suivants :

- la durée des échanges téléphoniques n'est pas prise en compte. En effet, les données sur la durée des échanges contiennent trop d'attrition pour pouvoir être exploitée.
- l'intensité des liens correspond au nombre total d'appels entre les individus totalisé sur une période de trois mois, ce qui correspond à une vague d'observation.
- la direction des appels entre les individus n'est pas prise en compte, afin de construire un réseau cohérent avec les choix explicités pour la construction du réseau « R_Friend » (cf. section 3.4.1).

Les liens sont par conséquent non orientés et leur intensité est la combinaison maximale du nombre d'appels orientés entre les paires d'individus.

Pour la collecte de données de contacts téléphoniques, nous avons identifié 112 individus qui apparaissent de façon permanente durant les 3 vagues d'observation. Par cette identification, nous nous assurons de l'absence de données manquantes en ce qui concerne les nœuds du réseau construit de la vague n°1 à la vague n°3, avec seulement 9 % de perte d'effectif sur le suivi de la communauté d'individus du réseau de référence « R_Friend ».

3.4.3. Le réseau uniplexe des co-présences « R_Prox »

Nous construisons le réseau des co-présences « R_Prox » à partir des relevés des proximités physiques quotidiennes entre les individus suivis. La période observée se déroule sur 3 vagues d'observation, de septembre 2010 à mai 2011, identiques à celles des contacts téléphoniques utilisés pour la construction du réseau « R_Call ».

La détection des proximités physiques quotidiennes entre les individus est réalisée par leurs téléphones mobiles grâce à un capteur de type « *Bluetooth* » intégré dans les appareils. Ainsi, toutes les 5 minutes une application logicielle détecte à partir de ce capteur si 2 ou plusieurs individus qui portent leur téléphone sur eux sont proches dans un rayon de 5 mètres environ (Eagle *et al.*, 2009). S'ils font partie des individus suivis par l'étude, les relevés de leurs identités sont enregistrés. Nous utilisons cet indicateur comme un « proxy » des co-présences et des interactions sociales synchrones en face-à-face entre les individus suivis. Le relevé est réalisé toutes les 5 minutes, tous les jours durant les 3 vagues d'observations de septembre 2010 à mai 2011.

Construction de l'indicateur de lien

Au sein de chacune des 3 périodes, nous exploitons la fréquence et le nombre quotidiens de détection des proximités physiques, sur une fenêtre horaire de 18h à 23h59. Afin de rendre compte des co-présences des individus suivis, nous construisons un indicateur calculé comme suit :

- la durée totale des co-présences entre 18h à 23h59 est totalisée pour chacune des 3 vagues d'observation.
- les liens reflètent la co-présence physique, ils ne sont donc pas orientés.
- l'intensité des liens correspond à la durée horaire de la co-présence physique des individus totalisée sur une période de 3 mois, ce qui correspond à une vague d'observation.

Pour la collecte des données des co-présences, 114 individus sont identifiés de façon permanente durant les 3 vagues d'observation. Par cette identification, nous nous assurons de l'absence de données manquantes en ce qui concerne les nœuds du réseau construit de la vague n°1 à la vague n°3, avec seulement 9 % de perte d'effectif sur le suivi des 123 individus de la communauté d'individus du réseau de référence « R_Friend ». Le tableau n°1.b synthétise la construction des 3 réseaux uniplexes.

Tableau n°1.b : description des 3 réseaux uniplexes formés à partir des données de l'expérimentation « Friends and Family »

Réseaux uniplexes	Nature des liens	Vagues d'observations	Nombre d'individus permanents pour toutes les vagues d'observation
R_Friend	proximité amicale liens non orientés	Septembre 2010 Décembre 2010 Mars 2011 Mai 2011	123
R_Call	contacts téléphoniques liens non orientés	Sep. 2010 - Nov. 2010 Déc. 2010 - Fév. 2011 Mars. 2011 - Mai. 2011	112
R_Prox	proximité physique liens non orientés	Sep. 2010 - Nov. 2010 Déc. 2010 - Fév. 2011 Mars. 2011 - Mai. 2011	114

3.4.4. Les réseaux multiplexes d'ordre 2 et 3

Pour la construction des réseaux multiplexes à partir des réseaux uniplexes, nous réalisons l'intersection (notée : « ∩ ») des effectifs des groupes d'individus toujours présents sur les 3 premières vagues d'observations pour chacun des réseaux. Nous indiquons, dans le tableau n°2, pour chaque configuration le nombre d'individus permanents sur toutes les vagues d'observations <u>et</u> communs aux réseaux uniplexes utilisés pour la formation des réseaux multiplexes.

Tableau n°2: réseaux multiplexes d'ordre 2 et 3 formés à partir des données de l'expérimentation « Friends and Family »

Réseaux multiplexes d'ordre 2 et 3	Vagues d'observations	Nombre d'individus communs et permanents sur les vagues d'observations
R_Friend∩R_Prox	Sep. 2010 - Nov. 2010 Déc. 2010 - Fév. 2011 Mars. 2011 - Mai. 2011	111
R_Friend∩R_Call	Sep. 2010 - Nov. 2010 Déc. 2010 - Fév. 2011 Mars. 2011 - Mai. 2011	109
R_Friend∩R_Prox∩R_Call	Sep. 2010 - Nov. 2010 Déc. 2010 - Fév. 2011 Mars. 2011 - Mai. 2011	104

Lecture : l'intersection des réseaux « R_Friend » et « R_Prox » forme un réseau multiplexe d'ordre 2, noté « R_Friend ∩ R_Prox », qui comprend 111 individus communs aux 2 réseaux uniplexes et permanents sur les 3 vagues d'observations considérées.

3.5. Modélisation longitudinale et dynamique des réseaux

Dans cette section, nous présentons la méthode mobilisée pour conduire les analyses longitudinales des 3 réseaux uniplexes - « R_Friend », « R_Call », « R_Prox » - et des 3 réseaux multiplexes. Dans une première partie, le modèle général utilisé simuler la dynamique temporelle des réseaux sociaux est présenté. Dans une deuxième partie, la spécification des modèles et des effets mobilisés susceptibles d'apporter des réponses à notre problématique est détaillée, ainsi que et la stratégie de modélisation suive.

Liens événementiels et liens durables

Pour les réseaux sociaux considérés, les liens qui les forment évoluent au sein d'une structure sociale susceptible de perdurer dans le temps d'observation, de la première à la troisième vague, soit 9 mois au cours desquels les étudiants du MIT étudient et vivent sur un même site. Si cela va de soi pour les liens formés par les enquêtes d'amitiés, il convient toutefois de le justifier pour les liens formés à partir des appels téléphoniques entre les individus et de leurs co-présences lors des temps libres.

En effet, ces indicateurs de communication relationnelle à distance pourraient être considérés comme restituant des « données événementielles » (« *event data* », Snijders *et al.* 2010, p.45) plutôt que des données d'états relationnels qui perdurent dans temps. Prosaïquement, par exemple, un appel téléphonique peut être adressé d'un individu à un autre pour solliciter une information (à quelle heure est programmée le cours demain ?), ou un service technique (as-tu le numéro de téléphone du plombier/baby sistter/uber-eats ?) sans que cela relève d'un lien d'amitié et de sociabilité consistant, durable dans le temps.

Pour surmonter ce potentiel écueil, comme nous l'avons décrit auparavant et suite aux préconisations de Snijders *et al.* (2010, p.45), les données issues des indicateurs d'appels téléphoniques et de co-présences ont été synthétisées sur chaque vague de 3 mois, afin de produire des observations de l'état d'une relation à chaque vague à partir de l'agrégation d'événements observés quotidiennement. De surcroît, nous appliquons un seuil sur les données afin de réduire ce que nous considérons comme du « bruit » dans le cadre de notre étude sur les relations amicales, c'est-à-dire pouvant relever d'une communication fonctionnelle ou technique (cf. chapitre 4). Ainsi, nous faisons l'hypothèse que les indicateurs agrégés des appels téléphoniques et des co-présences, tels que nous les avons construits avec les seuillages appliqués, restituent l'observation d'un lien social, créé, rompu ou maintenu, susceptible de durer dans le temps et qui ne se résume pas à un événement bref.

3.5.1. Le modèle SAOM - *Stochastic Actor-Oriented Models for network dynamics*

Le modèle général utilisé pour rendre compte de la dynamique temporelle des réseaux étudiés, est issu d'une famille de modèles nommés *Stochastic Actor-Oriented Models for network dynamics* (SAOM) (Snijders *et al.*, 2010) basée sur la théorie de l'action orientée par un but.

Ce type de modèle vise à élaborer une représentation de la dynamique d'évolution d'une structure relationnelle à partir d'au moins deux observations temporelles du réseau afin d'apprécier la significativité statistique de différents facteurs sur l'évolution de la structure. La structure est la variable indépendante, expliquée par des facteurs endogènes et exogènes. Le modèle SAOM permet d'élaborer une représentation, certes limitée mais fort utile, de la dépendance stochastique entre la création et la possible dissolution des liens qui forment la structure des réseaux. Cette modélisation permet de calculer la probabilité de changement de l'état (absent/présent) des liens établis par les acteurs du réseau.

L'ensemble de l'état des liens étudiés constituent l'état de la structure du réseau et il est fait l'hypothèse que les changements structuraux peuvent être interprétés comme un résultat d'un processus de Markov : l'état du réseau à chaque instant temporel t_i détermine la probabilité de son évolution ultérieure à l'instant t_{i+1}, sans d'autres effets provenant d'un état antérieur à t_{i-1}. Dans ce processus, les effets du passé du réseau sont considérés contenus dans les effets de son état présent (Snijders *et al.*, 2010). Cette modélisation s'appuie sur le cadre théorique de l'action de l'acteur orientée par un but. Ainsi, il est postulé que les acteurs du réseau (les individus qui sont reliés) choisissent d'établir, de maintenir, rompre des liens. Les choix des acteurs sont faits pour maximiser l'utilité du lien et, selon ce cadre théorique, à chaque temps t_i l'acteur est supposé contrôler l'établissement de liens avec d'autres acteurs et les conséquences de ces choix influencent la structure ultérieure du réseau qui constitue l'environnement relationnel des acteurs.

La simulation algorithmique SIENA - Simulation Investigation for Empirical Network Analysis

L'estimation des modèles est réalisée à l'aide de l'algorithme « *Simulation Investigation for Empirical Network Analysis* » (SIENA) développé par Snijders (2001). Cet algorithme itératif permet d'étudier la dynamique temporelle d'un réseau complet, dirigé ou non dirigé, observé au moins 2 fois dans le temps, pour un nombre fixé d'acteurs et sans prise en compte des éventuels poids des liens (seul l'absence ou la présence d'un lien est traitée). La modélisation résultante permet de tester des hypothèses explicatives concernant le changement longitudinal de réseaux complets, d'estimer la significativité statistique de facteurs et d'interactions entre facteurs, qui influencent l'évolution de la structure relationnelle d'un état observé à un instant t_i à un instant t_{i+1}.

Il est attendu que cette modélisation rende compte des dépendances statistique entre les observations constituées par les liens entre les acteurs du réseau. Par conséquent, les postulats classiques d'indépendance entre les observations ne sont pas respectés ici.

Postulats de la modélisation

Pour que cette modélisation puisse être opérationnalisée, plusieurs autres postulats sont posés. Ils concernent l'action temporelle des acteurs impliqués dans la structure relationnelle. Ils sont rassemblés en Annexes (p.167) selon la synthèse élaborée par Snijders *et al.* (2001, 2010) et de la Rúa (2004).

Certains de ces postulats nous apparaissent susceptibles d'être interrogés dans le cadre de notre étude. En particulier, les postulats selon lesquels chaque acteur a une connaissance totale de la structure relationnelle, qu'il n'a ni mémoire du passé, ni de capacité d'apprentissage, et qu'il est capable de déterminer les conséquences de ses choix dans l'établissement de liens d'amitiés, en termes de coûts et bénéfices et selon une stratégie de maximisation à court terme. La consistance de ces postulats est évaluée sous différents angles dans la littérature, notamment par Snijders (2010) et par la Rúa (2004).

En ce qui concerne le postulat de connaissance totale de la structure relationnelle par chaque acteur, Snjiders *et al.* (2010) soulignent qu'en pratique les acteurs n'ont besoin de disposer que d'une visibilité limitée de la structure relationnelle dans laquelle ils sont impliqué, avec un horizon délimité par une distance géodésique de valeur 2 (p.6). Le postulat de la rationalité limitée nécessite d'être particulièrement modulé dans le contexte de notre étude. En effet, les individus dont nous observons la structure relationnelle au sein du réseau formé par les liens d'amitié déclarée, sont des professionnels en reprise d'études. Ils forment, avec leurs co-résidents, des liens dont ils évaluent l'intensité. Ils révisent cette évaluation à 4 étapes (4 vagues d'observations).

Tel que nous l'avons argumenté dans notre chapitre introductif, nous considérons que ces individus, dans ce contexte spécifique, sont des acteurs « rationnels raisonnables » (Degenne et Forsé, 2004). Nous observons leurs comportements en adoptant le cadre d'un modèle de rationalité « élargie » (Forsé et Parodi, 2004) au sein duquel les acteurs comparent les avantages et les inconvénients d'une relation volontaire (l'amitié) et choisissent de nouer des liens selon les ressources recherchées pour atteindre des objectifs personnels, mais également professionnels (Coleman, 1988, 1990). Ainsi, dans notre étude, c'est le contexte qui étaye notre hypothèse selon laquelle l'acteur restitue lors des vagues d'enquêtes, la création, l'intensification, le maintien, la dévaluation ou la rupture d'un lien d'amitié.

En outre, si la modélisation du type SIENA ne permet pas de prendre en compte l'évolution nuancée de l'intensité des liens (ce type de modélisation ne traite que la présence ou l'absence d'un lien), nos données comprennent une graduation de l'intensité des liens : pour les liens d'amitiés, une échelle d'intensité de 7 niveaux discrets, pour les contacts téléphoniques et les co-présences, une échelle continue qui peut être discrétisée en catégories. Par conséquent, l'application de seuils sur l'intensité des liens nous permettent de nuancer et d'enrichir l'information expliquée par le modèle.

Toutefois, si la recherche d'objectifs personnels peut assez bien s'accorder avec le postulat selon lequel les acteurs forment des choix en suivant un stratégie de maximisation à court terme de l'utilité des liens établis, maintenus ou rompus, considérer comme nous le faisons dans cette étude que la poursuite de bénéfices est également d'ordre professionnel implique dès lors, de la part des acteurs, la poursuite d'une stratégie à plus long terme qui s'écarte du postulat d'une stratégie de maximisation immédiate. Par conséquent, nous nuancerons la force de ce postulat en reprenant à notre compte les observations émises par de la Rúa (2004) : il ne peut pas être réfuté l'existence de stratégies relationnelles à long terme (« je veux devenir ami de X car il est ami de Y », p.11), ni dans le champ de l'amitié ni, a fortiori, dans celui de la relation professionnelle. Il nous apparaît donc ainsi tout à fait pertinent de considérer que l'investissement relationnel de ces professionnels provenant du monde entier, en reprise d'études et pour un certain nombre loin de chez eux, qui partagent avec leur partenaire et leurs enfants une résidence étudiante pour un temps relativement court (9 mois), vise l'accès à des ressources instrumentales et pas seulement à des ressources expressives (Lin, 1995 ; Degenne et Forsé, 2004). Ainsi, les choix sont formés à cette fin selon des stratégies à moyen ou long terme.

Enfin, en ce qui concerne le postulat qui énonce que les acteurs n'ont ni mémoire du passé ni capacité d'apprentissage, il nous apparaît le plus fragile à soutenir dans le cadre de notre étude et, plus globalement, dans le champ des sciences humaines et sociales. Nous n'avons pas trouvé, dans la littérature sur la modélisation SAOM des réseaux sociaux, de propositions qui critiquent en particulier ce postulat. Ceci apparaît étonnant lorsque que les nœuds représentent des individus normalement doués de mémoire et de capacités cognitives assez largement partagées, qui leur permettent d'élaborer un apprentissage à partir de leurs expériences. N'est-ce pas une des bases fondamentales de l'évolution de l'espèce humaine ?

3.5.2. Spécification des modèles SIENA mobilisés

Dans cette section nous présentons la spécification des modèles SIENA, le choix des effets et du paramétrage. La variable dépendante des modèles est définie soit par un réseau uniplexe, soit par plusieurs réseaux qui constituent une structure multiplexe d'ordre 2 ou 3.

Les types de modèles

Chaque réseau non orienté est construit par un ensemble de *n* nœuds qui représentent *n* individus, reliés ou non entre eux par *k* liens, qui représentent, les relations amicales déclarées par les individus, les contacts téléphoniques ou les co-présences. Les variables constantes sont des co-variables monadiques : le sexe (0 – H ; 1 - F), l'origine ethnique (asiatique ou non, blanche ou non) et le nombre d'enfants (0, 1, 2 et plus) de chaque acteur durant la période d'observation.

Cinq types de modèles SIENA sont proposés pour les réseaux non orientés par Snijders et Pickup (2016) et un type de modèle pour les réseaux multiplexes (Snijders *et al.,* 2013). Leur formalisation dépend de la façon dont deux acteurs (*i* et *j*) agissent pour créer ou mettre fin à une relation. Quatre hypothèses sont formulées pour le processus de coordination entre les acteurs.

Premièrement, l'initiative peut être prise par l'un des deux acteurs (*one-sided initiative*) : l'acteur *i* a l'opportunité de réaliser un changement relationnel, la création ou l'extinction d'un lien s'il existe, et il fait ce choix unilatéralement. Deuxièmement, les deux acteurs *i* et *j* ont l'opportunité conjointe de faire un choix relationnel de création ou extinction d'un lien (*two-sided opportunity*). Dans ce cas, quatre processus sont alors envisagés (tableau n°3) :

1. un des deux acteurs impose unilatéralement son choix à l'autre pour la création d'un lien entre eux. Ce processus est représenté par le modèle nommé « ***forcing model*** ».
2. un des deux acteurs prend une initiative unilatérale et propose la création ou la dissolution d'un lien avec l'autre. Toutefois, cette fois-ci la création du lien est soumis à l'acceptation réciproque de cette initiative. S'il s'agit de la création d'un nouveau lien alors le second acteur doit confirmer. Sinon ce lien n'est pas créé. Pour la dissolution d'un lien, aucune confirmation n'est requise de la part du second acteur. Ce processus est exprimé par le modèle intitulé « ***initiative model*** ».
3. au moins un des deux acteurs décide de la création d'un lien, ce lien sera créé. Si les deux acteurs ne le souhaitent pas, le lien n'existera pas. Ce processus est traduit par le modèle nommé « ***pairwise forcing model*** ».
4. les deux acteurs décident de la création d'un lien entre eux sur la base d'un accord conjoint. Ce processus est exprimé par le modèle intitulé « ***pairwise joint model*** ».

Tableau n°3 : présentation synthétique des 4 types de modèles SIENA utilisés (d'après Snijders et Pickup, 2016; Ripley et al., 2020)

Modèle n°1 ***Forcing model***	Un acteur prend l'initiative et impose unilatéralement la création ou la dissolution d'un lien
Modèle n°2 ***Initiative model***	Un acteur prend l'initiative et propose la création ou la dissolution d'un lien. Si l'acteur propose un nouveau lien, l'autre acteur de la dyade doit confirmer, sinon le lien n'est pas créé. Pour la dissolution, la confirmation n'est pas requise.
Modèle n°3 ***Pairwise forcing model***	Une paire d'acteurs est choisie et l'existence d'un lien entre eux est reconsidérée. Le lien existera si au moins un deux acteurs choisit de créer un lien. Le lien n'existera pas si les deux ne le confirment pas.
Modèle n°4 ***Pairwise joint model***	Une paire d'acteurs est choisie et l'existence d'un lien entre eux est reconsidérée. Cela est basé sur la somme de leurs fonctions objectives pour l'existence de ce lien.

Dans nos modélisations, les modèles 1 et 2 donnent les mêmes résultats en ce qui concernent la significativité statistique et les coefficients des paramètres sont proches au 1/100ème. Ils sont ainsi présentés ensemble avec comme intitulé *« Forcing & Initiative model »*. Les modèles 3 et 4 (*Pairwise models)* sont présentés lorsqu'ils convergent et, le cas échéant, si le temps de calcul est inférieur à 48 heures.

Pour la modélisation des réseaux multiplexes, un seul modèle[8] est utilisé par l'algorithme SIENA (Snijders *et al.*, 2013 ; Ripley *et al.*, 2019-2020).

[8] Le modèle utilisé pour les modèles multiplexes n'est pas précisé dans le guide « RSiena » (Ripley *et al.*, 2020). Toutefois, à l'examen des scripts il semble que le modèle utilisé soit le n°1 « *Forcing model* ».

3.5.3. Construction des modèles

Nous utilisons dans cette étude une version simple du modèle SIENA avec une détermination basée uniquement sur la fonction objective qui exprime dans quelle mesure les différents changements de liens sont attractifs ou, formulé différemment, le niveau de satisfaction d'un acteur dans une configuration donnée du réseau. Au temps *t*, l'acteur a l'opportunité de modifier une relation : créer un lien, le maintenir ou le dissoudre. On admet comme postulat que l'acteur détermine son choix relationnel pour maximiser la valeur de la fonction objective (le niveau de satisfaction). Cette fonction comporte un élément aléatoire pour prendre en compte les raisons du choix de l'acteur qui ne sont pas exprimées par les effets étudiés.

Le paramétrage des modèles SIENA offre une large gamme d'effets possibles à considérer, exogènes et endogènes au réseau (Snijders *et al*., 2010 ; Snijders et Pickup, 2016 ; Ripley *et al*., 2020). Certains se révèlent assez ardus à interpréter (de l'avis même des auteurs). Nous faisons le choix de la parcimonie dans la sélection des effets, pour les effets endogènes structuraux et pour les effets exogènes relatifs aux attributs des individus. Notre choix est guidé par la préoccupation, d'une part, qu'ils contribuent à répondre à nos hypothèses et, d'autre part, de pouvoir être interprétés « sociologiquement » le plus aisément possible. Par ailleurs, dans le cas de réseaux non-orientés le choix des effets est plus limité que pour un réseau orienté et certains effets sont spécifiques comme nous le précisons dans la suite.

Les effets des variables socio-démographiques

Afin de répondre à notre hypothèse sur l'homophilie « H_homo » (cf. chapitre 3), nous incluons dans le paramétrage des modèles les effets relatifs aux variables socio-démographiques dont nous disposons au sein de la base de données mise à disposition par le MIT : le sexe des individus, leur origine ethnique et le nombre d'enfants des individus.

Le choix des effets liés à ces variables exogènes, constantes au cours de l'observation, doit nous permettre d'apprécier l'impact de l'homophilie sur la dynamique d'évolution des réseaux dans le temps. Ainsi, cela devrait contribuer à tester notre hypothèse « H_homo » d'une recherche évolutive des individus à nouer des liens avec d'autres individus semblables puis dissemblables. Nous incluons également 2 variables de contrôle que nous n'interprétons pas mais susceptibles de jouer un rôle comme vecteur de relations sociales : le nombre d'enfants et une interaction entre le sexe et le nombre d'enfants pour tenir compte du comportement connu genré lié à une charge familiale liée aux enfants.

Identité et similarité dyadique

Afin de tester l'hypothèse d'homophilie « H_homo », nous choisissons des effets susceptibles de répondre aux questions directrices suivantes :

- la sélection des relations est-elle réalisée selon une homophilie de sexe ou d'origine ethnique ?
- ou, au contraire, la significativité de ces effets révèle-t-elle d'une recherche de liens sociaux hétérophiles sur ces dimensions ?

De façon opérationnelle, nous choisissons un effet d'identité dyadique pour traduire le phénomène d'homophilie. Nous ajoutons un effet de similarité pour contrôler le rôle du nombre d'enfants comme vecteur facilitant les relations.

L'**effet d'identité dyadique** traduit que les individus qui partagent une caractéristique commune ont une probabilité plus élevée de créer et de maintenir un lien. Cet effet est construit sur l'égalité des valeurs prises par une même co-variable entre Ego et Alter :

- pour la variable catégorielle « sexe » binaire : Homme (0), Femme (1).
- pour la variable « ethnie », décomposée en 2 selon les origines ethniques les plus représentées :
    - ethnie asiatique (binaire : 0/1) : notée « ethnie_A »
    - ethnie blanche (binaire : 0/1) : notée « ethnie_B »

Cet effet est défini comme suit : $s_i(x) = \sum_j x_{ij}\, \mathrm{I}\{\vartheta_i = \vartheta_j\}$; où ϑ est la co-variable catégorielle considérée. La fonction $\mathrm{I}\{\vartheta_i = \vartheta_j\}$ = 1 si la condition $\{\vartheta_i = \vartheta_j\}$ est vérifiée, sinon $\mathrm{I}\{\vartheta_i = \vartheta_j\}$ = 0.

Cet effet est noté « ***même 'co-variable ϑ'*** » dans les tableaux de présentation des résultats des modélisations.

L'**effet de similarité** est utilisé pour contrôler les choix relationnels relatifs à une similarité du nombre d'enfants entre les individus, s'agissant ici d'une variable non dichotomique. La valeur de similarité est égale à 1 si la co-variable « nombre d'enfants » (codage ternaire : 0, 1, 2 enfants et plus) de deux individus a la même valeur, et 0 si l'un des deux a la plus haute valeur et l'autre la plus faible. La mesure étant centrée, la moyenne pour la valeur de la co-variable considérée est égale à 0. Cet indicateur permet de traiter cette variable catégorielle en maximisant les écarts entre deux acteurs ayant respectivement un seul enfant et 3 enfants ou plus.

L'effet de similarité d'une co-variable d'Ego et d'Alter est défini comme la somme centrée des valeurs de similarité d'une co-variable entre l'acteur (*i*) et les acteurs (*j*) avec lesquels l'acteur (*i*) est lié : $s_i(x) = \sum_j x_{ij}\,(sim_{ij}^{\vartheta} - \widehat{sim^{\vartheta}})$;

«ϑ» est la co-variable considérée et : $sim_{ij}^{\vartheta} = 1 - \frac{|\vartheta_i - \vartheta_j|}{\max x_{ij}|\vartheta_i - \vartheta_j|}$; avec $0 \leq sim_{ij}^{\vartheta} \leq 1$.

Cet effet est noté « ***similarité de la 'co-variable ϑ'*** » dans les tableaux de résultats des modélisations.

Les effets des co-variables d'Alter

Afin d'asseoir les résultats obtenus avec les effets relatifs à l'identité dyadique, il est recommandé dans le cadre de la modélisation SIENA d'inclure les effets de popularité pour les variables considérées. Ces effets apportent des réponses aux questions suivantes :

- le sexe, l'origine ethnique et le nombre d'enfants d'Alter sont-ils des attributs qui jouent un rôle statistiquement significatif observable sur l'évolution temporelle des relations ?
- y a-t-il des différences selon les réseaux observés ?

Afin d'apporter des éléments de réponse à ces questions, nous considérons tout d'abord **l'effet de popularité des co-variables** relative au sexe, à l'origine ethnique et au nombre d'enfants (*co-variate alter related popularity*). Il s'agit de d'apprécier ainsi l'effet de la popularité relatif à un attribut. Il s'agit de discerner si un individu qui possède un attribut spécifique est plus populaire dans les choix relationnels. Cet effet révèle l'importance de chacune des caractéristiques socio-démographiques dans l'évolution temporelle des relations au sein des réseaux observés. L'effet de la co-variable d'Alter est défini comme la somme des valeurs de la co-variable considérée pour tous les acteurs avec lesquels l'acteur (*i*) est lié, avec : $s_i(x) = \sum_j x_{ij}\vartheta_j$; où ϑ est la co-variable considérée. Cet effet est noté « ***'co-variable' d'Alter*** » dans les tableaux de résultats.

Finalement, nous incluons également comme élément de contrôle un **effet lié aux interactions** entre les 2 co-variables « sexe » et « nombre d'enfants ». Cet effet est noté « ***inter sexe * nbre d'enfants*** ».

Les effets structuraux « intra réseau »

Nous choisissons cinq effets structuraux susceptibles de contribuer à éclairer notre problématique et « sociologiquement » interprétables. Dans le cas d'un réseau non dirigé, la relation *i*→ *j* et *j*→*i* étant considérée comme équivalente elle est notée *i*—*j*. On note «x_{ij}» les liens relationnels du réseau pour un acteur (*i*) avec un acteur (*j*), avec *i* ≠*j*.

Tout d'abord, nous incluons **l'effet relatif à la densité** dans tous les modèles comme un effet de contrôle (Ripley *et al.*, 2020). Il exprime la tendance de base à nouer des liens. Tous les autres effets structuraux étant corrélés avec cet effet, ce paramètre ne sera pas interprété. L'effet de densité de degré est défini comme suit :

$s_i(x) = \sum_j x_{ij}$; où $\mathrm{x_{ij}} = 1$ s'il y a un lien entre i et j, sinon $\mathrm{x_{ij}} = 0$.

Cet effet est noté « ***degré (densité)*** » dans les tableaux de résultats des modélisations.

Ensuite, nous incluons dans le paramétrage des modèles, deux effets qui contribuent à l'exploration de nos **hypothèses structurales « hS1 » et « hS2 »** (cf. chapitre 1) d'une évolution du réseau qui traduirait une dynamique de recherche par les individus de nouer des liens avec des individus de plus faible connectivité au cours du temps : un effet d'appariement des individus et un effet relatif à la popularité et à l'activité des individus des réseaux.

Premièrement, nous considérons un effet susceptible de rendre compte de la tendance des individus à se lier avec des pairs structuralement dissimilaires en ce qui concerne le nombre de liens, la densité de leur voisinage. **L'effet d'appariement** (*degree assortativity*) nous apportera un éclairage sur l'appariement préférentiel des individus avec des pairs de degré équivalent ou non équivalent.

La significativité de ce facteur contribuera à discerner si des individus qui ont un degré élevé ont tendance à se lier avec des individus de degrés élevés ou avec des individus de degrés moins élevés et inversement. L'effet d'appariement des acteurs *i* et *j* est défini comme suit :

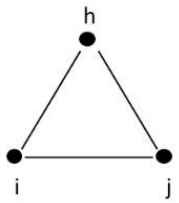
$s_i(x) = \sum_j x_{ij}\, x_{i+}^{1/c} x_{+j}^{1/c}$; avec : c = 2. Cet effet est noté « **appariement des individus de degré équivalent** » dans les tableaux de résultats.

Deuxièmement, nous incluons un **effet relatif à la popularité et à l'activité des individus** du réseau. D'une part, le paramètre de popularité permet d'apprécier le niveau de l'effet de « prestige », dû par exemple au niveau de nomination d'un individu dans les enquêtes sur les liens d'amitiés qui constituent la charpente du réseau « R_Friend ». D'autre part, le paramètre d'activité permet d'apprécier l'intensité de l'activité relationnelle. Un individu peut être observé comme très actif, par exemple par le nombre élevé de ses nominations d'amis, sans être toutefois lui-même populaire car peu nommé comme un ami par les autres individus.

Toutefois, dans le cas de réseaux non-orientés, où il n'y a pas de distinction entre les degrés intérieurs (liens entrants relatifs à une nomination, appels téléphoniques entrants) et les degrés extérieurs (liens sortants relatifs au fait de nommer un ami, appels téléphoniques sortants), l'appréciation de ces deux effets est traduite par la somme des effets d'activité et de popularité (*degree activity plus popularity effect*) (Snijders et Pickup, 2016 ; Ripley *et al.*, 2020). Les mêmes considérations sont faites pour les réseaux « R_Friend », « R_Call » et « R_Prox ».

L'effet relatif à la somme des effets de popularité et d'activité de l'acteur (*i*) est défini comme suit : $s_i(x) = \sum_j x_{ij}(x_{i+} + x_{j+})$. Cet effet est noté « **activité + popularité** » dans les tableaux de résultats.

Enfin, nous incluons deux effets relatifs aux configurations structurales triadiques pour répondre à notre hypothèse structurale « hS3 » (cf.chapitre 1). D'une part, un effet qui nous renseigne sur le nombre de triplets transitifs qui impliquent les individus (*transitives triades effect*). C'est un indicateur classique de la **transitivité** et de la **fermeture relationnelle** au sein d'un réseau social. Il permet d'opérationnaliser l'impact du phénomène relationnel : « les amis de mes amis deviennent mes amis » dans l'évolution du réseau. Cette configuration structurale est illustrée ci-dessous :

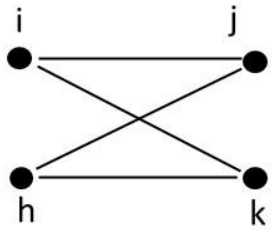


configuration structurale transitive
(réseau non-orienté)

L'effet du nombre de triades transitives qui impliquent l'acteur (*i*) est défini comme suit :

$s_i(x) = \sum_{j<h} x_{ij}\, x_{ih} x_{hj}$. Cet effet est noté « **triades transitives** » dans les tableaux des résultats des modélisations.

D'autre part, nous incluons un effet relatif au nombre de **configurations du type « 4-cycles »** (*number of four-cycles*) afin d'ouvrir l'appréciation de la transitivité/fermeture relationnelle du réseau à d'autres configurations que le nombre de triplets transitifs inclus précédemment. Cette configuration structurale est illustrée ci-dessous :



configuration structurale de type « 4-cycles »

(réseau non-orienté)

L'effet « nombre de cycle 4 » est formulé comme suit : $s_i(x) = \frac{1}{4}\sum_{j,k,h} x_{ij}\, x_{ik} x_{hj} x_{hk}$. Cet effet est noté « **4-cycles** » dans les tableaux des résultats des modélisations.

Les effets pour les réseaux multiplexes

L'évolution du réseau des relations amicales déclarées est susceptible de modifier le comportement des individus dans leurs interactions à distance et en présence et inversement. Considérés deux à deux (réseaux multiplexes d'ordre 2) ou tous ensemble (réseau multiplexe d'ordre 3), ils restituent une part de la multiplexité relationnelle, c'est-à-dire l'entrelacement des récits de la sociabilité des individus dans notre étude.

Afin d'apporter des réponses à nos hypothèses h_s4 et h_s5 sur l'entrelacement des relations, deux catégories d'effets multiplexes sont considérées : d'une part, **un effet dyadique principal** qui renseigne sur l'impact des choix relationnels des individus au sein d'un réseau sur leurs choix au sein d'un autre réseau, et, d'autre part, **trois effets triadiques mixtes** qui apportent un éclairage sur différents phénomènes de concordance et d'influence des choix relationnels d'un réseau à l'autre.

Dans la configuration de réseaux multiplexes d'ordre 2 et 3, pour chaque effet testé un des réseaux est considéré comme la variable dépendante (notée « X ») et l'autre réseau joue le rôle de variable explicative (notée « W »).

L'effet dyadique principal

L'effet dyadique que nous incluons dans le paramétrage des modèles multiplexes est l'effet principal d'un réseau sur l'autre (*dyadic entrainment effect)*. Il s'agit par exemple de l'effet direct du réseau W sur le réseau X. Cet effet de base contribue à apporter un éclairage sur l'influence d'un réseau sur l'autre et des éléments de réponses aux questions suivantes :

- comment les réseaux pris deux à deux co-évoluent-ils ?
- dans qu'elle mesure le fait d'échanger par téléphone (ou d'être en relation en présentiel) a-t-il, dans le temps, un impact sur l'évolution du réseau de nomination des amis ? Et inversement ?

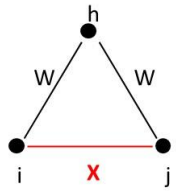
Notation : $s_i(x) = \sum_j x_{ij}\, w_{ij}$; avec : x_{ij} , les liens relationnels au sein du réseau X et w_{ij} les liens au sein du réseau W.

L'effet est noté « ***Network_X : Network_W*** ».

Les effets triadiques mixtes

Le premier des trois effets triadiques mixtes choisis doit nous apporter des éléments de réponses aux questions relatives à la concordance des choix relationnels des acteurs au sein du réseau X compte-tenu de leurs choix relationnels au sein du réseau W (*agreement along W leading to X)*. Ainsi, cet effet contribuera à apporter une réponse concernant les choix concordants ou discordants selon le mode de relation mobilisée et les déclarations des relations amicales.

Il s'agit par cet effet d'apprécier l'équivalence structurale, telle que représentée ci-dessous :



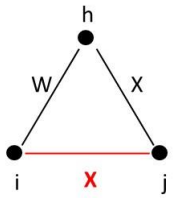
Agreement along W leading to X

la contribution de la relation $i \underline{^X} j$ au sein du réseau X est proportionnelle au nombre de choix relationnels communs au sein du réseau W : $i \underline{^W} h \underline{^W} j$.

Notation : $s_i(x) = \sum_{j \neq h} x_{ij}\, w_{ih}\, w_{jh}$; avec : x_{ij} , les liens au sein du réseau X, w_{ij} et w_{jh} les liens au sein du réseau W.

L'effet est noté « ***Network_X : from Network_W agreement*** ».

Un second effet triadique mixte est susceptible de nous renseigner sur la tendance des individus à faire les mêmes choix relationnels au sein du réseau X que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau W (*W leading to agreement along X*). Il s'agit ainsi d'apprécier l'effet de fermeture des chemins doubles mixtes entre les réseaux W et X (*closure of mixed W-X two-paths*), comme cela est figuré ci-dessous :



W leading to agreement along X

la contribution de la relation $i \underline{^X} j$ au sein du réseau X est proportionnelle au nombre de chemins mixtes doubles W-X : $i \underline{^W} h \underline{^X} j$.

Notation : $s_i(x) = \sum_{j \neq h} x_{ij}\, w_{ih}\, x_{hj}$; avec : x_{ij} et x_{hj} les liens au sein du réseau X, w_{ih} les liens au sein du réseau W. L'effet est noté « ***Network_X : Network_W to agreement*** ».

3.5.4. Stratégie de modélisation

Afin de tester l'hypothèse concernant l'homophilie et les effets des attributs socio-démographiques sur l'évolution des réseaux, les premiers modèles uniplexes sont construits sans facteurs structuraux, à l'exception du facteur relatif aux degrés qui traduit la densité des réseaux. Nous qualifions ces modèles de « naïfs » dans le sens où nous faisons « comme si » nous ignorions tous les autres effets structuraux (hormis celui général de la densité du réseau). Nous nous rapprochons ainsi des conditions d'une régression classique avec uniquement des effets propres aux variables socio-démographiques, du type régression logistique par exemple. Pour les modèles multiplexes, nous adoptons la même démarche, en incluant les facteurs aux effets inter-réseaux afin de prendre en compte l'impact de la multiplexité.

Convergence algorithmique et qualité d'ajustement des modèles

L'algorithme utilisé pour réaliser les modélisations est l'algorithme « RSiena » version 1.2-23 (Ripley *et al.*, 2020). Deux indicateurs permettent d'apprécier si l'algorithme d'ajustement du modèle a convergé de façon satisfaisante :

- le ratio global maximum de convergence (*overall maximum convergence ratio*)
- les statistiques *t* des écarts à la cible (*t statistics for deviations from targets*)

Les valeurs de référence, à considérer ensemble, proposées par Ripley *et al.* (2020) pour estimer la qualité de la convergence de l'algorithme d'ajustement du modèle sont présentées dans le tableau n°5 ci-dessous.

Tableau n°5 : valeurs de référence pour estimer la qualité de convergence de l'algorithme SIENA

overall maximum convergence ratio	< 0,25	bonne convergence
t statistics for deviations from targets	|t| < 0,1	
overall maximum convergence ratio	< 0,35	convergence proche
t statistics for deviations from targets	|t| < 0,15	

Selon Ripley et al. (2020, p. 21).

Dans notre étude, seuls les modèles qui satisfont aux deux caractéristiques suivantes sont retenus, présentés et analysés : convergence des « *t ratios* » < 0,1 et « ratio global maximum de convergence » < 0,25. Les seuils de significativité statistiques considérés dans cette étude sont les suivants : « † : p < 0,1; * : p < 0,05; ** : p < 0,01; *** : p < 0,001 ».

Similarité des réseaux observés successivement

L'index dit de Jaccard (1900), mentionné dans les tableaux qui décrivent les changements dans les liens entre les périodes d'observations, est un index de similarité entre les réseaux observés successivement. Cet index permet d'apprécier la stabilité. Il est calculé comme suit pour 2 vagues d'observations successives de l'état d'un réseau (Snijders *et al.*, 2010) : $\frac{N11}{N11+N01+N10}$

Avec N_{ij} : le nombre de liens avec la valeur *i* dans une vague d'observation et la valeur *j* pour la vague suivante. Ainsi, N_{11} est le nombre de liens présents dans chacune des vagues d'observation du réseau, N_{01} le nombre de liens nouveaux et N_{10} le nombre de liens dissous. Selon Sjniders *et al.* (2010) et Ripley *et al.* (2020), si les index de Jaccard sont très bas alors que le degré moyen ne progresse pas fortement, cela signifie que le « *turn-over* » est peut-être trop élevé pour considérer que les données relationnelles forment un réseau en évolution. Dans ce cas, une modélisation de type SIENA n'est pas appropriée.

Pour des réseaux dont les degrés moyens changent de valeur entre 2 et 15 selon les vagues d'observation, des index de Jaccard de 0,3 et plus sont considérés comme corrects (Snijders *et al.*, 2010, p.49). Des valeurs inférieures à 0,2 signalent que l'estimation sera difficile à conduire. Toutefois, si les index sont faibles parce que le réseau est en croissance avec une création de nouveaux liens, ou en décroissance avec une extinction de liens, la modélisation de type SIENA doit normalement aboutir. De même, des réseaux clairsemés (avec la plupart des degrés inférieurs à 2) peuvent conduire à des valeurs d'index de Jaccard faibles, sans poser de problème pour la modélisation SIENA (Ripley *et al.*, 2020, p.20).

Qualité d'ajustement des modèles (goodness of fit)

La qualité d'ajustement des modèles peut être appréciée par plusieurs statistiques, telles que la distribution des distances géodésiques et/ou la distribution des degrés par exemple. Nous adopterons ces deux statistiques pour apprécier différemment la qualité des modèles obtenus. Un algorithme développé par Lospinoso (2012) permet d'estimer la qualité d'ajustement, selon la statistique considérée, selon l'amplitude des écarts entre les valeurs observées et les valeurs simulées par les modélisations SIENA à la fin de chaque période. Un test basé sur la distance de Mahalanobis est utilisé pour calculer une p-value qui permet avec une représentation graphique des écarts d'apprécier la qualité de l'ajustement (Ripley *et al.*, 2020). Lorsque l'une des deux statistiques est médiocre, un test est réalisé par l'addition de variables temporelles muettes aux vagues d'observation 1 à n-1, afin de vérifier que ceci n'est pas dû à une hétérogénéité temporelle (Lospinoso *et al.*, 2011).

Enfin, nous choisissons de paramétrer les algorithmes au niveau le plus exigeant en ce qui concerne le nombre d'itérations. Bien que cela augmente les temps de calcul, ceci permet d'accroître la stabilité de la détermination algorithmique itérative des écarts-types (Lospinoso et Snijders, 2015).

Ressources calculatoires

Le matériel informatique disponible pour exécuter les algorithmes de modélisation et les temps de calcul nécessaires pour aboutir à des convergences acceptables sont des éléments à ne pas négliger en analyse de réseaux. Afin d'aboutir aux résultats des modélisations présentées ci-après, il a été nécessaire de mobiliser le matériel suivant :

- 2 ordinateurs mobiles « i3 Core Intel Pentium » avec 8 Go de mémoire vive,
- 1 ordinateur mobile « i5 Core Intel Pentium » avec 16 Go de mémoire vive,
- 1 serveur professionnel disposant d'une mémoire vive de plus 64 Go, mis à disposition par la Très Grande Infrastructure de Recherche « TGIR Huma-Num » (CNRS, Université d'Aix-Marseille et Campus Condorcet).

Ainsi, par exemple, un des modèles incluant 7 effets et exécuté sur le serveur de la plateforme en ligne « TGIR Huma-Num », a conduit à un temps de calcul de 48 heures avant de livrer des résultats.

S'il est difficile d'anticiper précisément ces incontournables aspects temporels et matériels en amont, plusieurs essais exploratoires nous ont conduit à respecter une certaine sobriété dans le paramétrage des modèles afin de produire des résultats dans un temps raisonnable, c'est à dire au maximum 50 heures pour la simulation algorithmique d'un modèle. Pour information, nous avons simulé et exploré au total plus de 30 modèles.

4. Résultats

Dans ce chapitre, nous présentons les résultats des traitements des 3 bases de données relationnelles utilisées pour répondre à nos hypothèses à notre problématique générale. Ces résultats sont obtenus à partir des 3 réseaux uniplexes non-orientés, des deux réseaux non-orientés multiplexes d'ordre 2 et finalement du réseau non-orienté multiplexe d'ordre 3 « R_Friend∩Call∩Prox ». Nous concentrons nos analyses sur les résultats relatifs à l'origine ethnique tout en contrôlant les attributs relatifs au sexe et au nombre d'enfants lors des modélisations.

Introduction

Au chapitre 3 nous avons présenté l'élaboration de 3 réseaux uniplexes de nature différente en ce qui concerne la relation sociale entre les individus :

- le réseau « R_Friend » restitue des déclarations de relations d'amitiés.
- le réseau « R_Call » restitue des contacts factuels à distance.
- le réseau « R_Prox » restitue des contacts factuels en présence.

Notre étude portant sur l'évolution des liens d'amitié, le réseau de référence est celui qui restitue l'évolution des relations amicales déclarées (« R_Friend »).

Nos réseaux sociaux sont observés à l'issue d'intervalles temporels de 3 mois chacun, que nous désignons à l'aide du terme de « vague ». Pour chacune des vagues d'observations, une analyse descriptive des 3 réseaux nous permet de saisir les caractéristiques fondamentales de chacun d'entre eux à partir d'un ensemble de métriques de connectivité issues de la théorie des graphes et de la sociologie des réseaux sociaux (section 4.2).

En ce qui concerne l'analyse statistique inférentielle, nous considérons qu'à l'issue des vagues d'observations, les liens créés, rompus et maintenus, sont le résultat, d'une part, des positions structurales occupées par les acteurs du réseau et, d'autre part, de caractéristiques individuelles et de similarités qui caractérisent une homophilie socio-démographique ainsi que d'un résidu dû aux phénomènes aléatoires non expliqués par les éléments modélisés (section 4.3). Ainsi, l'analyse des 3 réseaux uniplexes prend en compte les dimensions suivantes :

- d'une part, l'évolution du réseau résulte de facteurs issus de la structure elle-même. Il s'agit d'effets structuraux endogènes au réseau observé, tels que la densité relationnelle, la transitivité des relations, l'appariement préférentiel des acteurs selon leur connectivité.
- d'autre part, l'évolution temporelle de la structure est influencée par des facteurs exogènes : des co-variables monadiques constantes dans le temps de l'observation. Il s'agit des attributs socio-démographiques des acteurs : le sexe, l'origine ethnique et le nombre d'enfants.

Présentation socio-démographique de la communauté suivie

Les attributs socio-démographiques des 123 individus suivis, sont : le genre, l'origine ethnique[9] des individus ainsi que le nombre d'enfants déclarés (tableau n°4) :

- le sexe des individus suivis : Homme/Femme.
- le nombre d'enfants déclarés : de 0 à 4.
- l'origine ethnique[10] des individus suivis : asiatique, noire, hispanique, israélienne, moyen-orientale, métissée, amérindienne, blanche.

Tableau n°4 : description socio-démographique de la communauté des étudiants suivis

Sexe	**Homme**	**Femme**	**Total**
%	50,4	49,6	100

Nombre d'enfants	**0**	**1**	**2 et plus**	**Total**
%	58,5	24,5	17	100

Origine ethnique	**Asiatique**	**Noire**	**Hispanique**	**Israélienne**
Nombre d'individus	47	3	11	1
%	38,5	2,5	9	0,8
Origine ethnique	**Moyen-Orientale**	**Métissée**	**Amérindienne**	**Blanche**
Nombre d'individus	7	2	1	50
%	5,7	1,6	0,8	41

Effectif : 123 individus. Une donnée manquante pour l'origine ethnique.

La communauté suivie est équilibrée en genre et environ 40% des individus suivis ont un ou plusieurs enfants. En ce qui concerne l'origine ethnique déclarée par les individus, la communauté apparaît « multi-ethnique » dans son ensemble, ce qui correspond à l'attractivité mondiale attendue pour le MIT qui est la 4ème université mondiale dans le classement de « Changaï » pour l'année 2020[11]. Cependant, une polarisation très nette apparaît avec 2 groupes d'individus qui dominent avec une quasi-égalité d'effectif : des étudiants d'origine ethnique blanche et asiatique qui représentent ensemble près de 80% de la communauté, soit 97 étudiants.

[9] Dans l'expérimentation « *Friends and Family* » (Aharony *et al*., 2011), l'origine ethnique des personnes est mentionnée par le terme « *race-ethnicity* ». Nous choisissons le terme « origine ethnique ».

[10] Les catégories mentionnées dans la base de données sont les suivantes : « *Asian, Black, Hispanic, Israely, Middle Eastern, Mixed-race, Native American, White* »

[11] Le classement de « Shanghai » réalisé par l'« *Academic Ranking of World Universities* » désigne depuis 2003 les 1 000 meilleures universités au monde. Données publiées le 15/08/2020 (www.shanghairanking.com).

La troisième catégorie en effectif est celle des étudiants « hispaniques » (11 individus), suivie des étudiants du « moyen-orient » (7), ainsi que 3 étudiants « noirs », 2 étudiants qui se déclarent d'origine « métissée » et finalement un étudiant « israélien » et un « amérindien ». Un étudiant n'a pas donnée d'information sur son origine ethnique.

Personnalité des étudiants suivis

En début de suivi longitudinal, les individus sont invités à répondre à un questionnaire avec plusieurs assertions qui portent sur leur personnalité et la représentation de soi. Les réponses doivent être apportées aux assertions suivantes sur une échelle de 1 (pas d'accord) à 7 (tout à fait d'accord) :

- Assertion n°1 : « Je recherche activement autant d'informations que possible dans une nouvelle situation ».
- Assertion n°2 : « Je recherche de nouvelles opportunités pour grandir en tant que personne (par exemple : de l'information, les ressources d'autrui) ».
- Assertion n°3 : « Partout où je vais, je recherche de nouvelles choses ou expériences ».

Les réponses de 117 des individus indiquent qu'environ les 3/4 des étudiants suivis, lorsqu'ils sont placés dans une nouvelle situation, se déclarent très actifs dans la recherche d'informations, de ressources détenues par leurs pairs, de nouvelles expériences. La communauté suivie est ainsi dominée par des individus qui se présentent comme ayant une démarche très ouverte aux opportunités et expériences que peuvent offrir les rencontres nouvelles. Les graphiques qui présentent la distribution des réponses sont proposés en Annexes (graphiques A.1, p.160).

4.1. Analyses descriptives des réseaux uniplexes

4.1.1. Le réseau « R_Friend »

Le réseau « R_Friend » non orienté est bâti à partir des liens d'amitié déclarés entre les 123 individus, via un questionnaire auto-administré (cf. section 3). Le questionnaire porte sur l'intensité des liens d'amitiés entre les individus sur une échelle de 0 (aucun lien d'amitié) à 7 (forte proximité amicale). Il est reconduit à l'identique 4 fois : en septembre 2010 (vague n°1), décembre 2010 (vague n°2), mars 2011 (vague n°3) et mai 2011 (vague n°4).

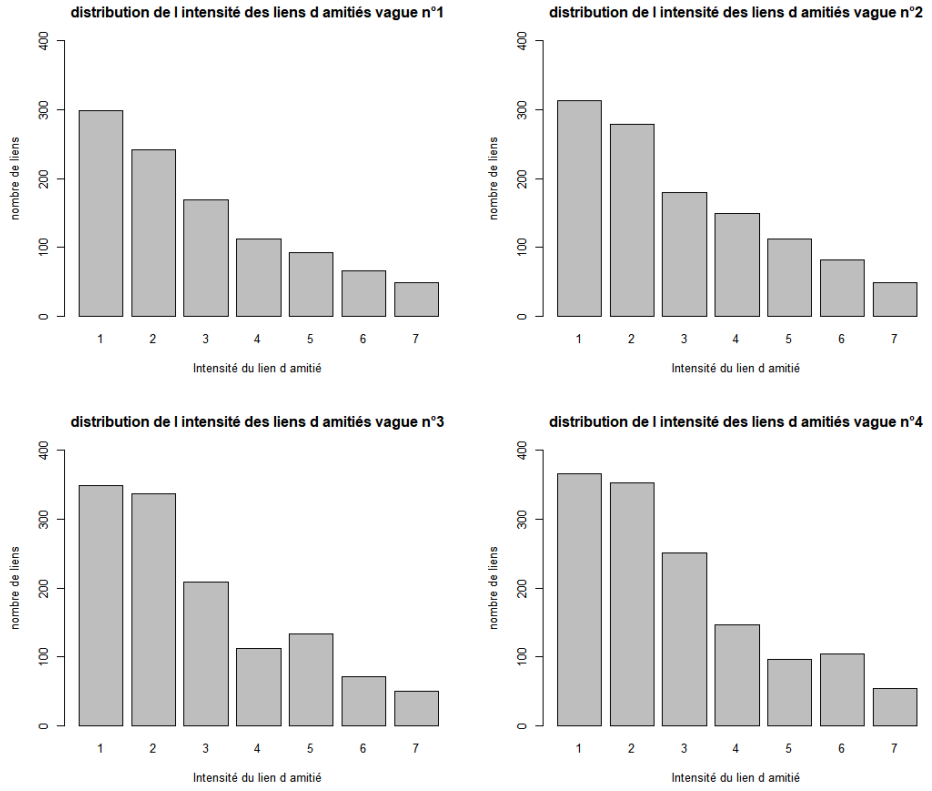
Métriques du réseau des liens d'amitiés déclarés « R_Friend »

Le tableau n°5 présente différentes métriques de cohésion, de transitivité, de centralité et leurs évolutions pour le réseau « R_Friend » (sans et avec un seuil sur l'intensité des liens) par vagues. La distribution de l'intensité des relations amicales déclarées est présentée sur le graphique n°1 et la distribution des degrés des nœuds du réseaux est présentée sur le graphique n°2.

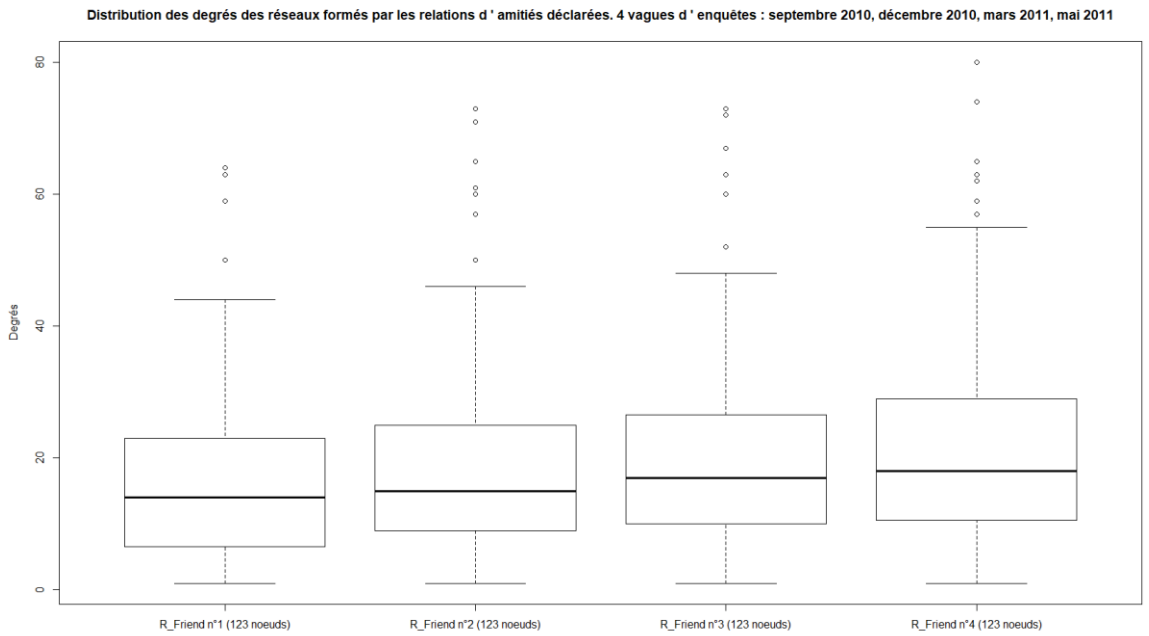
Tableau n°5 : principales métriques du réseau « R_Friend » et évolution selon les vagues d'observation

Réseau « R_Friend » non dirigé Liens sans seuil sur l'intensité *[seuil sur l'intensité des liens : ≥3]*	**Vague n°1**	**Vague n°2**	**Vague n°3**	**Vague n°4**
Taille du réseau	123 nœuds, 1030 liens *[490]*	123 nœuds, 1162 liens *[571]*	123 nœuds, 1258 liens *[574]*	123 nœuds, 1370 liens *[653]*
Densité	0,137 *[0,065]*	0,154 *[0,076]*	0,167 *[0,076]*	0,182 *[0,087]*
Longueur moyenne du chemin le plus court *(Average Path Length)*	2,15 *[2,83]*	2,01 *[2,78]*	2,02 *[2,74]*	1,99 *[2,59]*
Coefficient global d'agglomération *(Global Clustering Coefficient)*	0,41 *[0,428]*	0,41 *[0,424]*	0,43 *[0,415]*	0,46 [*0,394]*
Coefficient local moyen d'agglomération *(Local Clustering coefficient)*	0,59 *[0,535]*	0,55 *[0,522]*	0,55 *[0,533]*	0,59 *[0,524]*
Centralisation de Freeman normalisée	0,39 *[0,188]*	0,44 *[0,186]*	0,43 *[0,202]*	0,47 *[0,232]*
Centralité de degré normalisée				
Moyenne	0,14 *[0,065]*	0,15 *[0,076]*	0,17 *[0,076]*	0,18 *[0,087]*
Ecart-type	0,11 *[0,058]*	0,12 *[0,064]*	0,12 *[0,064]*	0,14 *[0,070]*
Degré maximal	0,52 *[0,254]*	0,60 *[0,260]*	0,60 *[0,278]*	0,65 *[0,319]*
Centralité d'intermédiarité normalisée - valeur maximale *(standardized betweenness centrality)*	0,35 *[0,208]*	0,34 *[0,200]*	0,15 *[0,130]*	0,12 *[0,100]*
Origine ethnique des 10 premiers étudiants intermédiaires clefs (*gatekeepers)*	Blanche (7), Hispanique (1), Asiatique (1), inconnue (1)	Blanche (6), Hispanique (1), Asiatique (1), Moyen-Orientale (1) inconnue (1)	Blanche (7), Hispanique (1), Asiatique (1), inconnue (1)	Blanche (6), Métissée (1), Asiatique (2), inconnue (1)

Graphique n°1 : distribution de l'intensité des liens d'amitiés du réseau « R_Friend »



Graphique n°2 : distribution des degrés des nœuds du réseau « R_Friend ».



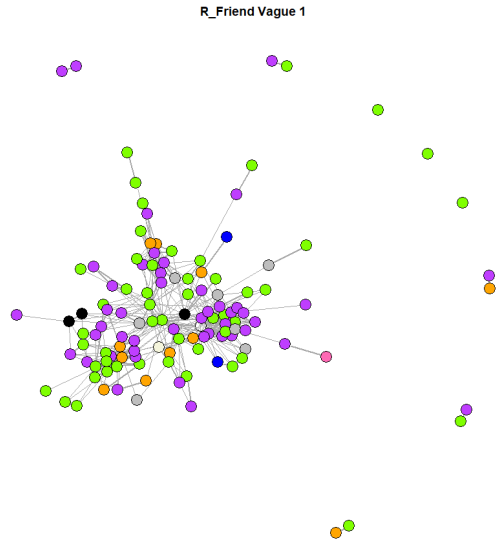
Effectif : 123 individus. Nombre de liens : vague n°1 : 1030; vague n°2 : 1162; vague n°3 :1258; vague n°4 : 1370.

Le graphique n°3 représente l'évolution du réseau avec un seuil sur l'intensité des liens d'amitié de 3. Ce seuil correspond à une relation amicale hiérarchisée selon les catégories suivantes :

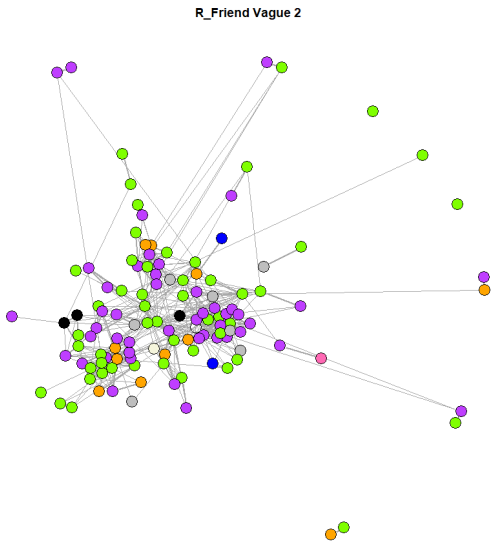
- *3: « This person is a friend (low ranking) » ; 4: « This person is a friend (medium ranking) » ;*
- *5: « This person is a friend (high ranking) » ; 6: « This person is a close friend » ;*
- *7: « This person is family or as close to me as a family member ».*

Sur le graphique n°2, l'épaisseur des liens est proportionnelle à l'intensité de la relation. La couleur des nœuds révèle les origines ethniques des individus. L'algorithme utilisé pour le tracé du graphe *(layout algorithm*) est celui développé par Fruchterman et Reingold (1991) et les coordonnées des nœuds sont fixées sur ceux du premier graphe.

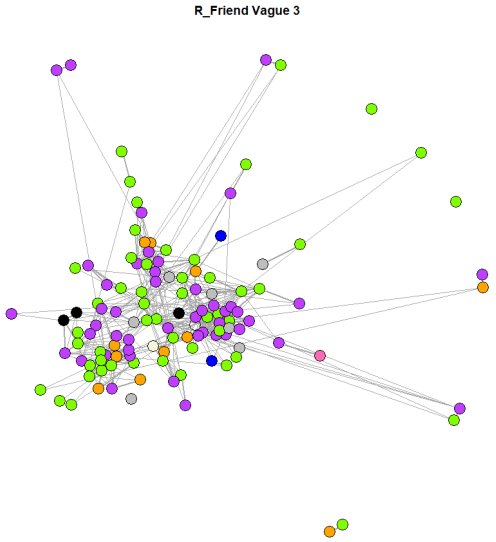
Graphique n°3 : évolution du réseau « R_Friend » (liens d'intensité ≥ 3)



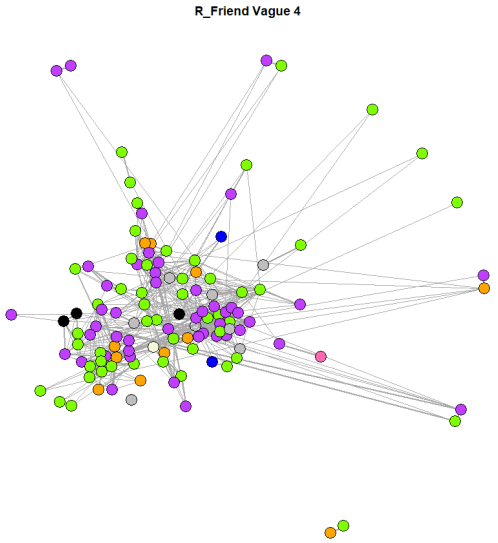
Vague n°1 : 123 nœuds, 490 liens, intensité 3 à 7



Vague n°2 : 123 nœuds, 571 liens, intensité 3 à 7



Vague n°3 : 123 nœuds, 574 liens, intensité 3 à 7



Vague n°4 : 123 nœuds, 653 liens, intensité 3 à 7

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

La cohésion globale du réseau « R_Friend » croît de façon strictement monotone au fil du temps : la densité du réseau « R_Friend » progresse de 25% entre la première et la dernière vague (tableau n°5), avec ou sans l'application d'un seuil sur l'intensité des liens. Ainsi, le nombre de liens d'amitiés déclarés augmente entre les individus de la communauté. A l'examen du graphique n°3 (réseau seuillé), quelques individus apparaissent non connectés à un pair : ils sont situés en périphérie, sur les graphes de la vague n°1 (3 individus), n°2 (2 individus) et n°3 (2 individus) et tous d'origine ethnique asiatique. Quelques binômes, peut-être des couples, se situent également en périphérie sans autre lien que celui qui les unit. A la dernière vague, tous les individus sont connectés à au moins un autre et un seul binôme apparaît entretenir une relation exclusive.

Ces indicateurs nous permettent d'apprécier que la communauté observée longitudinalement d'un point de vue structural est dynamique en ce qui relève de la création de liens d'amitiés déclarés, avec une nette tendance à la densification de sa cohésion globale.

A l'exclusion des individus non connectés, la distance moyenne la plus courte entre deux nœuds (*Average Path Length*, tableau n°5) décroît de façon monotone, avec une réduction de 8% (sans seuil) ou 7% (avec seuil) de la distance moyenne entre la première et la dernière vague d'observation. En moyenne, 2 individus sont connectés par 2 intermédiaires dans leurs relations amicales (hors individus non connectés). Le nombre moyen d'intermédiaires est un peu plus élevé pour le réseau seuillé (2,74). Pour ce réseau des relations d'amitiés affirmées (seuil de 3), si l'on observe ainsi une décroissance de la distance moyenne entre les individus au fil du temps elle demeure plus élevée qu'au sein du réseau global.

Les valeurs d'agglomération globale (*Global Clustering Coefficient*, tableau n°5) se situent dans l'écart [0,3; 0,6] souvent constaté pour les réseaux empiriques. Pour le réseau « R_Friend » sans seuil, la probabilité moyenne que deux nœuds soient connectés sachant qu'ils ont un voisin un commun est de 43%, avec une augmentation de 5% entre la première vague et la dernière. En revanche, pour le réseau seuillé les valeurs chutent de 7% entre la première et la dernière vague. Par ailleurs, au sein du réseau des relations amicales affirmées (seuil de 3) la transitivité relationnelle locale est relativement stable alors qu'elle décroît au sein du réseau global qui comprend des liens d'intensité faibles et fortes.

L'évolution de l'indicateur de centralité de degré normalisée suggère que l'importance (du point de vue de la centralité) de certains individus augmente au fil du temps : les valeurs moyennes et maximales sont croissantes, avec un degré maximal de 0,65 (0,32 avec seuil) atteint lors de la dernière vague. La progression des valeurs centralité de degré normalisée est encore plus importante (en valeur relative) pour le réseau seuillé.

Ainsi, au sein du réseaux d'amis (seuil de 3), le potentiel d'interactions directes augmente de façon plus forte qu'au sein du réseau global, ce qui est un résultat attendu. L'origine ethnique des 5 étudiants les plus actifs dans la connectivité est par ordre décroissant de degré : asiatique, hispanique et blanc[12].

L'indicateur de centralisation (Freeman) croît de façon quasi-monotone dans le temps : **plus le temps passe, plus le réseau apparaît dominé, structuralement, par un nombre faible d'individus**, au sein du réseau global (sans seuil) et au sein du réseau des relations d'amitiés affirmées (seuil de 3). Parmi ces individus, certains se distinguent par la valeur élevée de la centralité d'intermédiarité normalisée. Ces valeurs élevées distinguent les acteurs qui connectent des groupes différents et occupent une position d'intermédiaires clefs (***gatekeepers***). Comme cela est aisément visible sur le graphique n°4 (p.94) où le seuil appliqué sur l'intensité des liens est de niveau 5 (ce qui correspond à la réponse : « *This person is a friend (high ranking)* », ils forment parfois le seul « pont » entre des groupes. Pour le réseau non seuillé, de façon peu évolutive de la première à la dernière vague, les 10 plus importantes valeurs sont détenues très largement par des étudiants d'origine ethnique blanche (70 %), puis de façon plus réduite, des étudiants d'origine ethnique asiatique (12,5 %), hispanique (7,5 %) et moyen-orientale (2,5 %). Pour le réseau avec un seuil de 3 pour l'intensité des liens, les valeurs sont les suivantes : étudiants d'origine ethnique blanche (65 %), asiatique (35 %), hispanique (5 %) et moyen-orientale (5 %).

Pour le réseau non seuillé, la perception du rôle fort d'intermédiaire pour les individus d'origine ethnique blanche dans le réseau des relations d'amitiés déclarées est nette. En effet, ces derniers occupent les 3/4 des places des premiers intermédiaires clefs alors qu'ils ne représentent que 41% des individus de la communauté. Inversement, les étudiants d'origine ethnique asiatique, dont l'effectif au sein de la communauté (38,5 %) est proche de celui des étudiants d'origine ethnique blanche, comptent 5 fois moins d'individus « *gatekeepers* » que ces derniers. En outre, les étudiants d'origine ethnique asiatique comptent à peu près autant d'intermédiaires clefs que les étudiants d'origine ethnique hispanique dont l'effectif au sein de la communauté est 4 fois plus faible (9 %).

Pour le réseau des relations d'amitiés (seuillé à 3), le constat énoncé ci-dessus est plus nuancé mais garde sa pertinence : les individus d'origine ethnique blanche rassemblent 50% des premiers « *gatekeepers* » et les étudiants d'origine ethnique asiatique en comptent 35%.

[12] Par malchance (pour cette étude) un étudiant d'origine ethnique « non renseignée » a le degré le plus élevé.

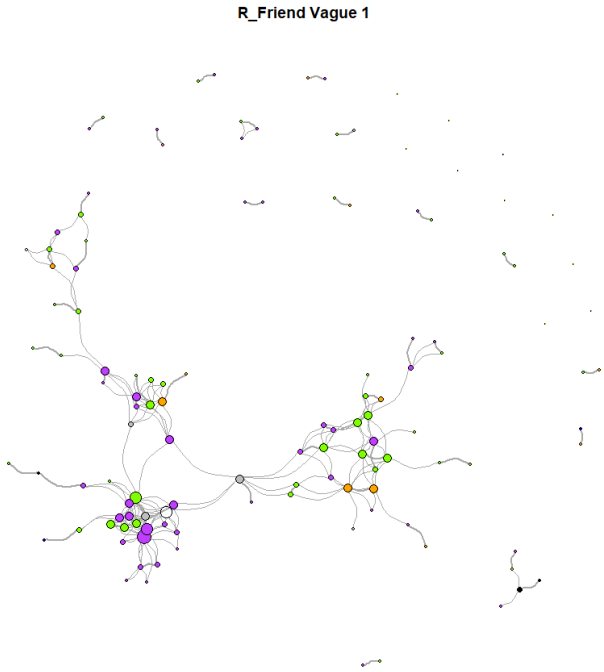
L'évolution des valeurs maximales de l'indicateur de centralité d'intermédiarité normalisée apportent un éclairage sur les individus qui occupent une cette position structurale particulière. En effet, les valeurs décroissent nettement à partir de la troisième vague. **Ainsi, les individus positionnés comme les premiers intermédiaires durant les deux premières vagues perdent à partir du 6ème mois de vie collective, de façon nette, leur importance relative en tant que « connecteurs » de groupes.**

Les graphes suivants (graphique n°4, p.94) représentent l'évolution du réseau avec un seuil d'intensité de 5, afin de mettre en valeur les relations amicales fortement affirmées (« *This person is a friend (high ranking)* »). Sur ces graphes, l'épaisseur des liens est proportionnelle à l'intensité de la relation et la taille des nœuds est proportionnelle aux degrés des individus. Les graphes font clairement apparaître entre 3 et 4 grappes principales selon les vagues. Afin de confirmer cette observation, nous mettons en œuvre 2 algorithmes de détection de possibles communautés au sein des graphes avec une seuil de 5 pour l'intensité de pour le lien d'amitié.

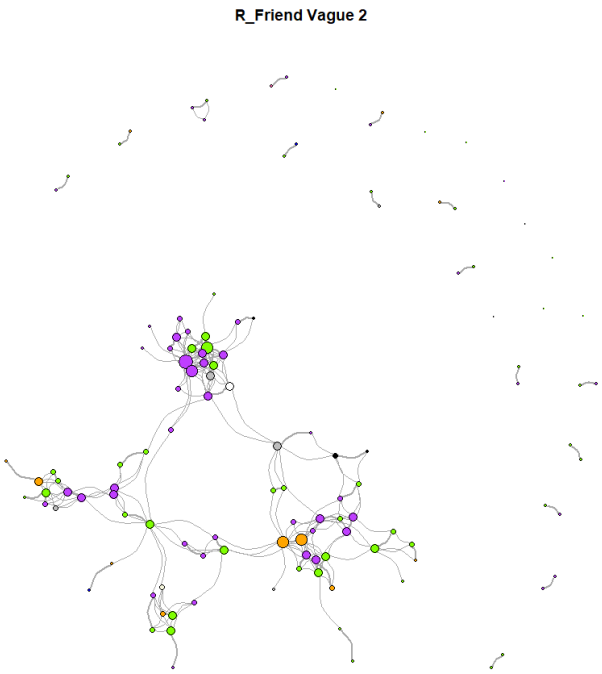
Le premier algorithme « *cluster edge betweenness* » (Newman et Girvan, 2004) est basé sur l'indicateur d'intermédiarité des nœuds. La centralité d'intermédiarité (*betweenness centrality*) dénombre le nombre de fois où un nœud est positionné comme un point de passage intermédiaire le long du plus court chemin entre deux autres nœuds.

Le second algorithme « *cluster fast greedy* » (Clauset *et al.*, 2004) est construit à partir d'une optimisation du score de modularité, c'est-à-dire une mesure de la qualité du partitionnement des nœuds du réseau en communautés. Il fait apparaître les parties du graphe les plus connectées entre elles, avec une forte densité interne et une faible densité externe. Le graphique n°5 (p.95) représente des résultats de détection de communautés possibles obtenus par l'algorithme « *cluster edge betweenness* », pour la vague n°1 et la vague n°4. Sur ces graphes, l'épaisseur des liens est proportionnelle à l'intensité des liens d'amitié.

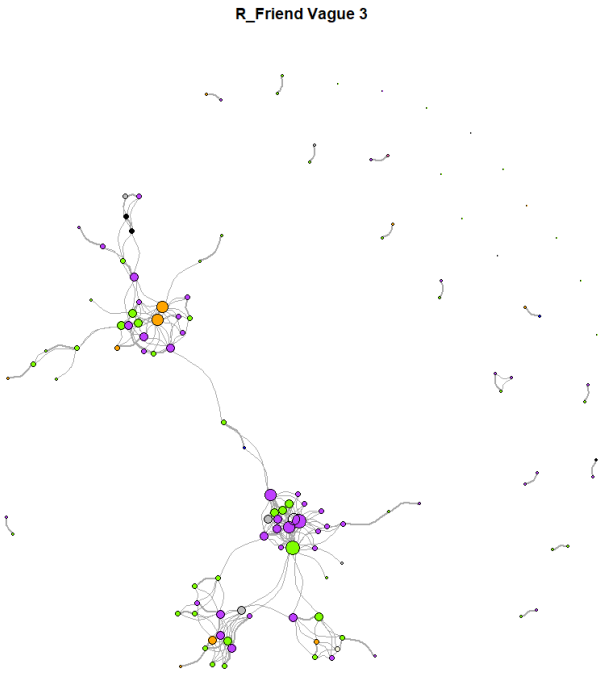
Graphique n°4[13] *: représentation de l'évolution du réseau « R_Friend » pour les 4 vagues d'observation*



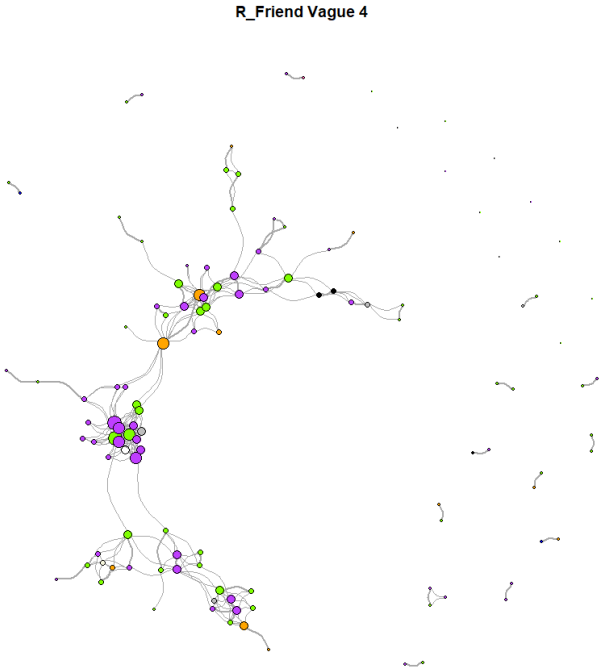
Vague n°1 : 123 nœuds, 208 liens, intensité : 5 à 7



Vague n°2 : 123 nœuds, 243 liens, intensité : 5 à 7



Vague n°3 : 123 nœuds, 254 liens, intensité : 5 à 7



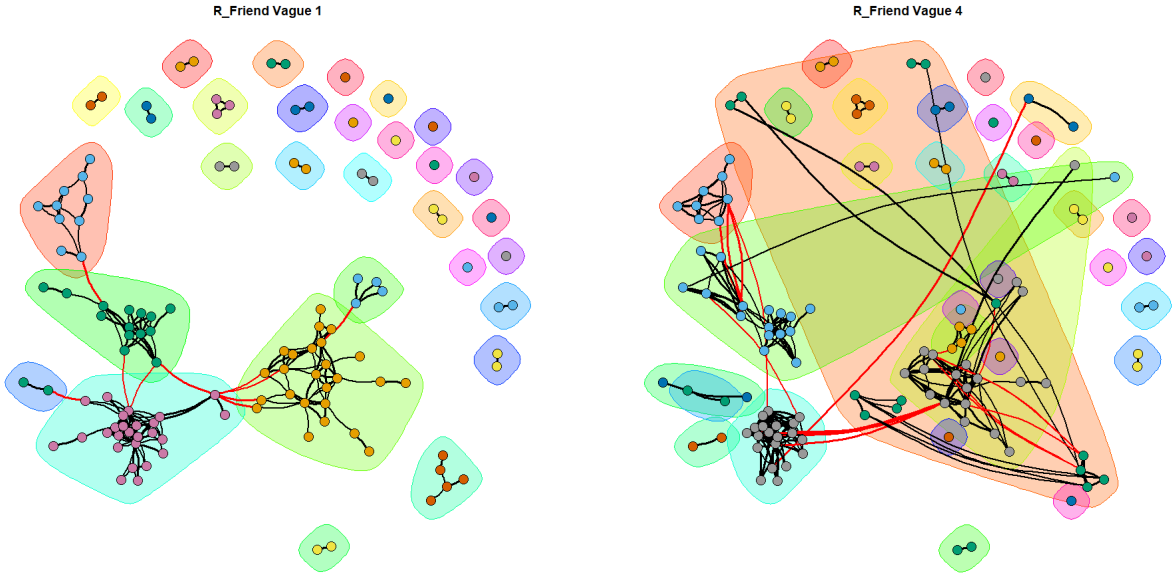
Vague n°4 : 123 nœuds, 255 liens, intensité : 5 à 7

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

[13] Afin d'accroître la lisibilité du tracé, sur le graphique n°4 les coordonnées des nœuds des vagues n°2 et suivantes ne sont pas fixées sur ceux du premier graphe. Algorithme de tracé des graphes : Fruchterman et Reingold (1991).

Les résultats obtenus à l'issue des détections algorithmiques de communautés confirment et précisent l'observation précédemment énoncée. Lors de la première vague, les algorithmes mettent en évidence 4 communautés principales ainsi qu'une quarantaine d'individus plus périphériques. Les 2 principaux groupes rassemblent chacun 20% des individus, le 3ème et 4ème groupe en regroupent chacun 10% environ. Les individus périphériques non connectés au 4 principaux groupes représentent 30% de l'ensemble. L'évolution au cours du temps montre l'apparition de liens, parfois denses, entre les principaux groupes et surtout avec des individus périphériques. Ainsi, à la vague n°4 seulement 20% des individus ne sont pas connectés à un des 4 principaux groupes.

Graphique n°5 : détection de communautés au sein des graphes (algorithme : « cluster edge betweenness », Newman et Girvan, 2004). Intensité des liens : 5 à 7. Les couleurs sont associées aux communautés.



modularité : 0,71 modularité : 0,73

L'examen des graphes (graphique n°4 et n°5) montrent notamment 2 groupes principaux :

D'une part, un groupe qui comprend en majorité des individus d'ethnie « Asiatique » avec pour certains d'entre eux des degrés très élevés, ainsi que quelques individus d'origine ethnique blanche avec des degrés élevés. Les 6 autres ethnies sont peu ou pas représentées dans ce groupe. Le nombre de liens nouveaux créés au fil du temps avec les autres groupes et les individus en périphérie apparaît faiblement croissant.

D'autre part, un groupe qui comprend, en majorité et à effectifs proches, des individus d'ethnique asiatique et blanche avec des degrés plus élevés pour ces derniers. Par ailleurs, dans ce groupe, on observe quelques individus d'origine ethnique hispanique dont les degrés sont de plus en plus élevés au fil du temps. Ce groupe apparaît tisser un nombre de liens croissant avec de plus petits groupes ou des individus périphériques isolés.

Ces deux groupes sont connectés par un seul individu « *gatekepper* » d'origine ethnique blanche (vague n°1) puis hispanique (vague n°4).

Synthèse

La cohésion du réseau des liens d'amitiés déclarées augmente au fil des vagues d'observation, pour aboutir en dernière vague à ce qu'aucun individu ne soit isolé. La dynamique de connectivité est forte, le nombre de liens est croissant au sein du réseau global et du réseau des relations d'amitié affirmées (seuil de 3 : « *This person is a friend (medium ranking)* »). Néanmoins, nous observons que la transitivité globale chute au fil du temps au sein du réseau des relations d'amitiés affirmées, contrairement à l'évolution de la transitivité locale qui demeure relativement constante. En outre, la distance moyenne entre 2 individus est plus faible au sein du réseau global qu'au sein de celui des relations d'amitiés affirmées. Dans ce dernier, il apparaît une stratification relationnelle plus stable qu'au sein du réseau dans son ensemble : les amitiés affirmées restent localisées et stables dans leur configuration transitive au cours du temps contrairement aux transitivités relationnelles considérées pour la globalité du réseau.

Par ailleurs, les réseaux des relations d'amitiés (globales et affirmées) sont de plus en plus dominés par un nombre faible d'individus, notamment d'origines ethniques asiatiques, hispaniques et blanches. Ces individus ont un rôle important en termes de connectivité : ils ont une activité dominante et occupent les positions d'intermédiaires clefs (*gatekeepers*). Cette domination décroit en force et se concentre sur quelques-uns au fil du temps. Enfin, nous observons que les étudiants d'origines ethniques blanches et asiatiques, tout en ayant des effectifs équivalents, concentrent de façon non équilibrée les positions structurales clefs.

4.1.2. Le réseau « R_Call »

Cent-douze individus sont identifiés comme permanents pour les 3 vagues des échanges téléphoniques vocaux : de septembre 2010 à novembre 2010 (vague n°1), de décembre 2010 à février 2011 (vague n°2) et de mars 2011 à mai 2011 (vague n°3). L'effectif compte moins de 10% d'attrition (11 individus) en comparaison avec le réseau « R_Friend ». Les individus qui constituent ce réseau forment un noyau stable d'enquêtés, c'est-à-dire un groupe d'individus que l'on retrouve de façon permanente comme répondants à toutes les vagues. Les liens non orientés sont les contacts téléphoniques entre les individus. L'intensité des liens est fonction du nombre quotidien d'appels totalisés pour chacune des vagues d'observation. Le tableau n°6 présente les métriques du réseau.

Tableau n°6 : métriques du réseau « R_Call » pour les 3 vagues d'observation.

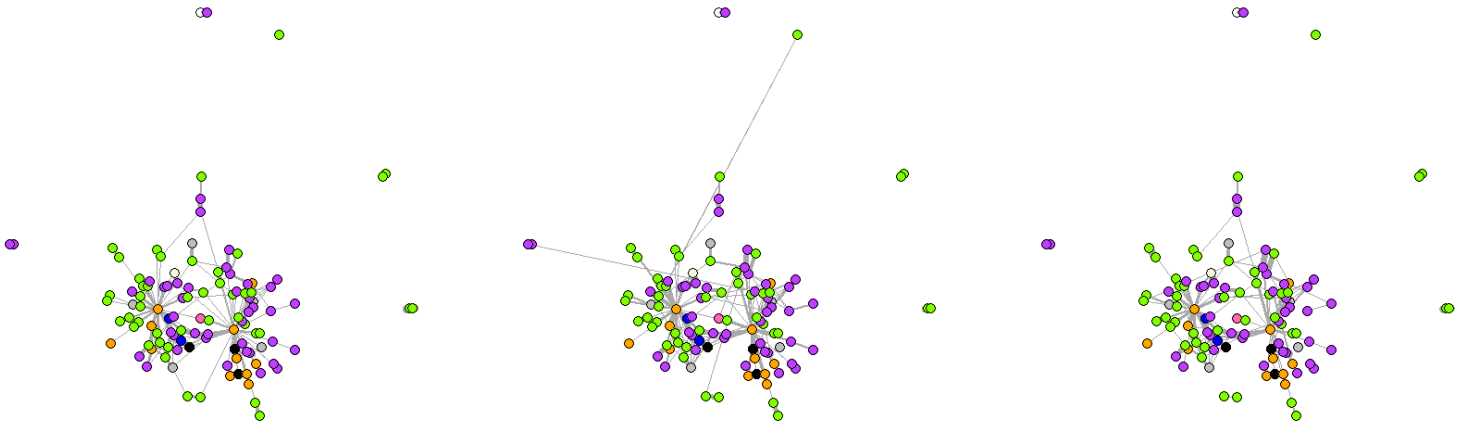
Réseau « R_Call » non dirigé intensité des liens : seuil de 3 appels minimum par vague	**Vague n°1**	**Vague n°2**	**Vague n°3**
Taille du réseau	112 nœuds, 190 liens	112 nœuds, 200 liens	112 nœuds, 183 liens
Densité	0,030	0,032	0,029
Longueur moyenne du chemin le plus court *(Average Path Length)*	3,252	3,508	3,441
Groupements connexes *(Clusters*[14]*)* nombre [taille]	6 « clusters » [2, 103, 1, 2, 2, 2]	6 « clusters » [2, 104, 2, 1, 2, 1]	12 « clusters » [2, 94, 1, 2, 2, 2, 1, 2, 2, 2, 1, 1]
Coefficient global d'agglomération *Global Clustering Coefficient*	0,130	0,167	0,195
Coefficient local moyen d'agglomération *Local Clustering Coefficient*	0,461	0,461	0,390
Centralisation de Freeman normalisée	0,365	0,301	0,231
Centralité de degré normalisée			
Moyenne	0,030	0,032	0,029
Ecart-type	0,044	0,041	0,036
Degré maximal	0,396	0,333	0,261
Centralité d'intermédiarité normalisée - valeur maximale *(standardized betweenness centrality*)	0,529	0,477	0,381
Origine ethnique des 10 premiers étudiants « *gatekeepers* »	Hispanique (4), Asiatique (3), Blanche (3)	Hispanique (3), Asiatique (3), Blanche (3), Moyen-Orientale (1)	Hispanique (3), Asiatique (3), Blanche (4)
Modularité *(*algorithme « *cluster edge betweenness* »*)*	0,75	0,74	0,64

[14] Groupements détectés par l'algorithme « *Clusters* » (Csardi et Nepasz, 2005 ; boîte à outil « Igraph » du logiciel R).

Représentation du réseau « R_Call »

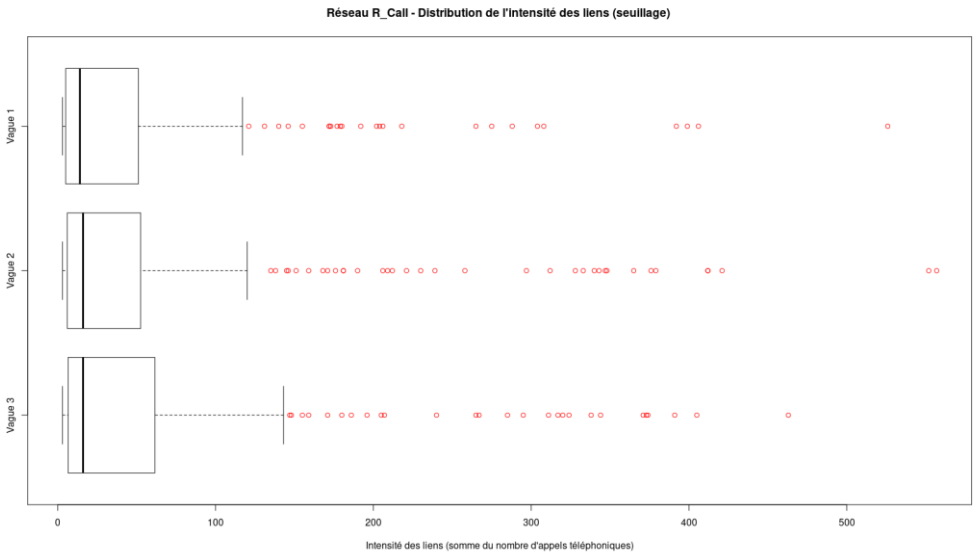
Les graphiques n°6 et n°8 illustrent les échanges téléphoniques entre les individus sous la forme de graphes non dirigés. L'algorithme utilisé pour le tracé des graphes *(layout algorithm)* est celui développé par Fruchterman et Reingold (1991). Les coordonnées des nœuds des vagues n°1 à n°3 sont fixées sur ceux du premier graphe. Sur le graphique n°8, la taille des nœuds est proportionnelle degrés (nombre de contacts téléphoniques). Nous faisons l'hypothèse qu'un certain nombre d'appels ne concerne pas un échange substantif du registre amical, mais correspondent à une communication fonctionnelle ou technique, ce que nous considérons comme du « bruit » dans le cadre de notre étude sur les relations amicales. Après un examen des données (statistiques descriptives présentées en Annexes (tableau A3, p.164), nous appliquons un seuillage au 1er quartile du nombre total d'appels par vague afin de réduire ce bruit. Ce seuil correspond à un appel par mois au minimum, soit 3 appels par vague.

Graphique n°6 : représentation de l'évolution du réseau « R_Call » (seuil : 3 appels/vague)



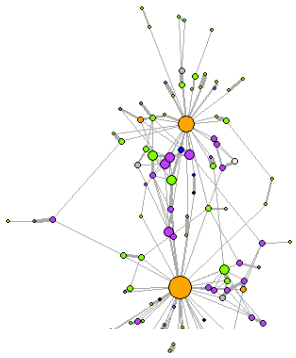
Vague 1 : 112 noeuds, 190 liens Vague 2 : 112 noeuds, 200 liens Vague 3 : 112 noeuds, 183 liens

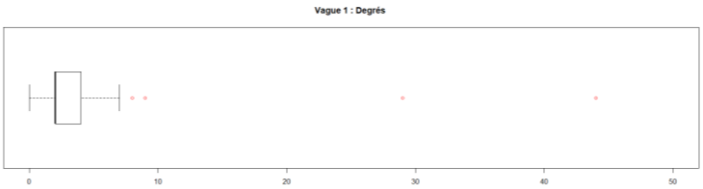
Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

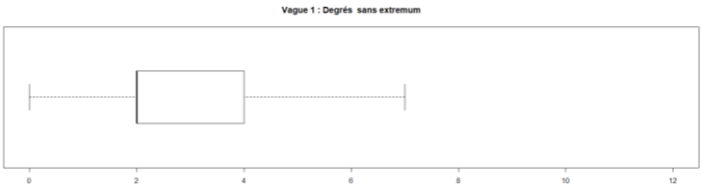


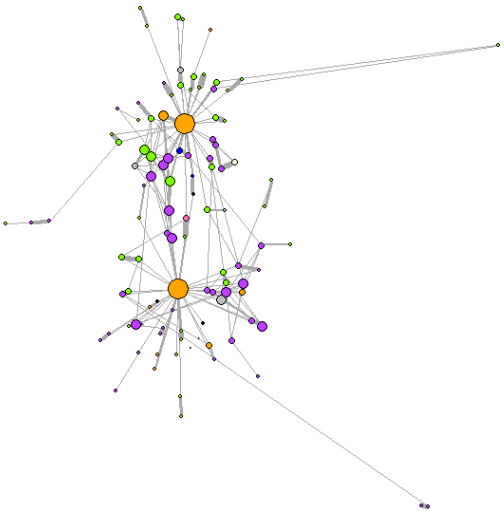
Graphique n°7 : distribution de l'intensité des liens du réseau « R_Call » (seuil : 3 appels/vague)

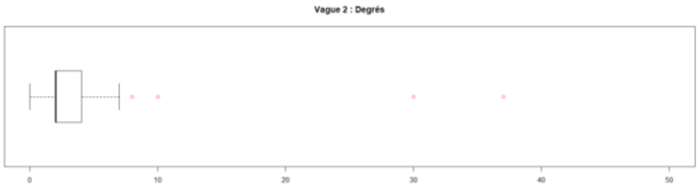
Graphique n°8 : représentation de l'évolution du réseau « R_Call » (seuil) et distribution des degrés.

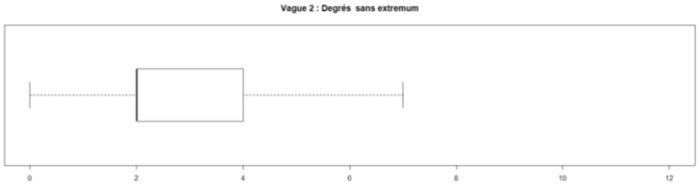


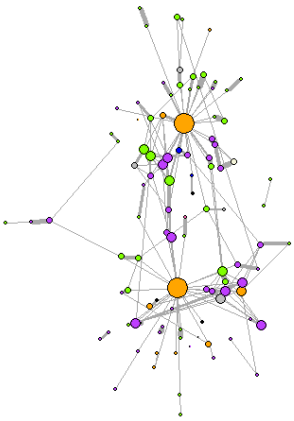


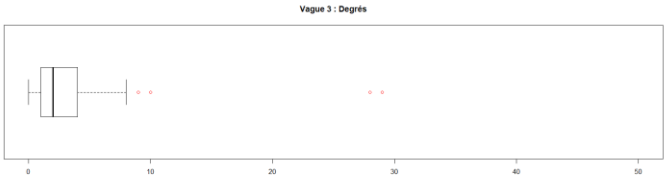


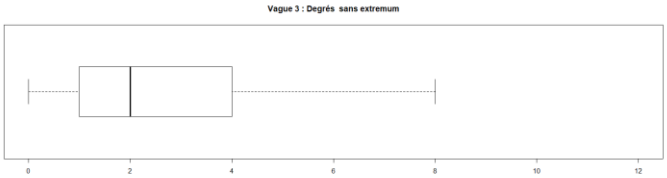












Vague 1 : 112 noeuds, 190 liens ; Vague 2 : 112 noeuds, 200 liens ; Vague 3 : 112 noeuds, 183 liens.

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

La densité du réseau « R_Call » est quasi-constante au fil du temps, avec une valeur moyenne très faible (0,03) en comparaison avec celle du réseau « R_Friend » (0,16 en moyenne) (tableau n°6). Le nombre de liens par contacts téléphoniques évolue peu de la première à la dernière vague et la centralité de degré normalisée moyenne est stable dans le temps avec, toutefois, une chute assez nette de la valeur du degré maximal lors de la dernière vague d'observation du réseau. Plus globalement, les valeurs maximales des degrés du réseau « R_Call » sont environ 2 fois plus faibles que celles observées pour le réseau « R_Friend ». Par ailleurs, contrairement au réseau « R_Friend », le réseau « R_Call » tend à se décentraliser avec des valeurs de centralisation de Freeman normalisées décroissantes au fil du temps.

L'examen de la distribution de l'intensité des liens (graphique n°7) révèle qu'une trentaine d'individus a des contacts téléphoniques d'une fréquence très élevée avec certains de leurs pairs, de 7 à 10 fois supérieure à la valeur moyenne observée de 5 contacts par semaine. Ces valeurs extrêmes doivent probablement correspondre aux échanges inter-couples. Plus globalement, les individus se répartissent dans l'intensité de leurs contacts autour de la valeur médiane d'un appel par semaine (Annexe, tableau A.3). L'intensité des contacts téléphoniques croît de plus de 25% de la première à la deuxième vague d'observation, ce qui correspond à une période de constitution de la communauté et de primo-découverte des pairs. Ensuite, l'intensité des contacts à distance reflue de 14% entre la deuxième et la dernière vague.

L'examen du graphique n°6 permet de grossièrement discerner **2 amas principaux d'individus, l'un qui semble plutôt investi par des étudiants d'origine ethnique blanche** (au Nord des graphes) **et l'autre par des étudiants d'origine asiatique** (au Sud des graphes)**, avec toutefois une mixité relativement faible mais perceptible au sein des 2 agglomérats majeurs.** Au fil du temps, les 2 principaux amas semblent relativement figés en ce qui concerne leur composition ethnique. L'examen des indicateurs de groupement (tableau n°6) révèlent qu'une dizaine d'individus, exclusivement d'origine asiatique et blanche, forment des foyers (*clusters*) de faibles tailles (de 6 à 12 individus) en périphérie du groupement principal (une centaine d'individus).

En ce qui concerne la transitivité observée, le réseau « R_Call » traduit une dynamique de connectivité sensiblement différente de celle des relations de proximité amicale déclarées. D'une part, le coefficient d'agglomération global les valeurs sont très inférieures à celles constatées pour le réseau « R_Friend », mais elles progressent nettement, avec une augmentation de plus 30% entre la première et la dernière vague d'observation, alors qu'elles étaient quasiment constantes pour le réseau des amitiés déclarées. Ainsi, au fil du temps on observe une augmentation effective de la transitivé globale.

D’autre part, pour le réseau de échanges par téléphone, les valeurs des coefficients locaux moyens d’agglomération sont très nettement supérieures à celles des coefficients globaux moyens. Par conséquent, la transitivité locale est nettement plus importante que la transitivité globale. Ce contraste est plus marqué qu’au sein du réseau des relations d’amitiés déclarées. En outre, la longueur moyenne du chemin le plus court est nettement plus importante dans le réseau des échanges téléphoniques (3,4) que dans celui des relations amicales (2). Par conséquent, les relations d’échanges par téléphone apparaissent plus « cloisonnées » que pour le réseau « R_Friend ».

Les graphes seuillés à 3 appels au minimum par vague (graphique n°8, p.99) révèlent 2 individus d’origine ethnique hispanique qui se distinguent par leurs degrés très élevés, 10 fois supérieurs au degré moyen de l’ensemble de la communauté suivie. L’examen des 10 premières valeurs maximales de l’indicateur de centralité d’intermédiarité normalisée met en évidence le rôle proéminent d’étudiants d’origine ethnique hispanique, blanche et asiatique dans la connexion entre les 2 principaux agglomérats d’individus. Les trois origines ethniques - blanches, asiatiques et hispaniques - disposent du même nombre d’individus intermédiaires (*gatekeepers*), alors que la dernière (hispanique) représente seulement 9% des individus de la communauté et les deux premières chacune environ 40 %. En outre, de la première à la dernière observation du réseau « R_Call », 2 individus d’origine ethnique hispanique apparaissent tout à la fois très actifs et en position de « *gatekeepers* ». Ils relient les 2 principaux amas dont la composition ethnique est plutôt dominée d’une part, par des étudiants d’origine blanche et, d’autre part, par des étudiants d’origine asiatique. Cependant, l’évolution des valeurs maximales de centralité d’intermédiarité montre que ce rôle clef dans les échanges relationnels perd de l’importance au fil du temps, ce qui semble converger avec l’augmentation observée de la transitivé globale : les individus se contactent de plus en plus directement entre eux et de moins en moins *via* un intermédiaire.

Synthèse

Les métriques descriptives du réseau « R_Call » révèle un récit de la sociabilité dont l’intensité effective des relations entre les individus qui contraste avec l’intensité déclarée par le réseau « R_Friend ». Le nombre de contacts téléphoniques entre les individus est relativement constants dans le temps. Une fermeture relationnelle locale apparaît et le réseau se décentralise au cours du temps. Deux groupes plutôt homogènes se différencient selon les 2 origines ethniques majoritaires. Des individus d’origine hispanique, minoritaires au sein de la communauté, jouent un rôle majeur de connecteurs entre ces 2 groupes. Enfin, des petits groupes d’individus, exclusivement d’origine asiatique et blanche, forment des foyers de faibles tailles en périphérie des groupements principaux.

4.1.3. Le réseau « R_Prox »

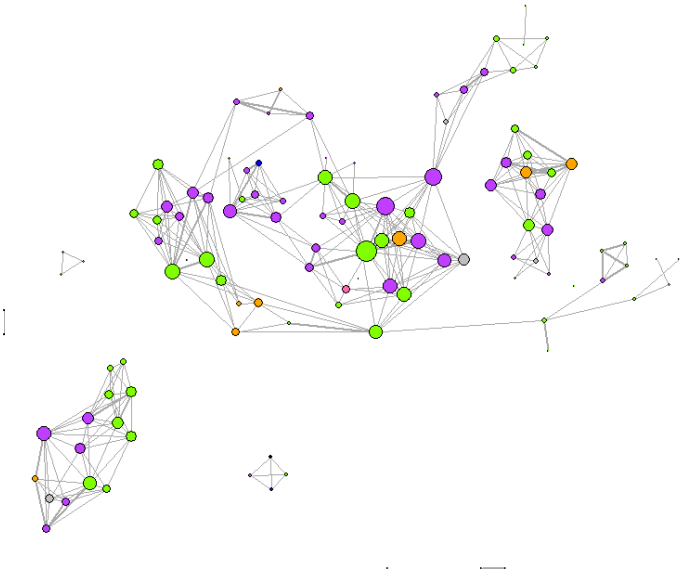
Pour les données concernant les contacts en proximité physique, 114 individus sont identifiés comme permanents pour les 3 vagues, d'une durée de 3 mois chacune. Ces individus forment un noyau stable d'enquêtés, c'est-à-dire un groupe d'individus que l'on retrouve de façon permanente à toutes les vagues. Comme nous l'avons précisé au chapitre 3, l'indicateur utilisé est calculé par la somme du nombre d'heures quotidiennes de co-présence physique des individus de 18h à 23h59.

Nous opérons un seuillage des liens basé sur un nombre d'heures minimum de co-présences détectées[15]. Ceci nous permet de filtrer les échanges qui résulteraient de rencontres fortuites, de très faibles durées, des situations de co-présences aléatoires brèves qui correspondent putativement aux croisements des individus au sein de la résidence étudiante familiale ou à l'extérieur (campus et lieux proches) sans générer un échange en face-à-face dialogué consistant. Nous considérons ces brèves rencontres comme une forme de « bruit » dans les données pour lequel nous faisons l'hypothèse que cela ne s'inscrit pas dans une relation dialoguée d'amitié. Toutefois, nous veillons à ne pas supprimer totalement les faibles temps de co-présence. En effet, ceux-ci peuvent être du registre des rituels de croisement au sein de la résidence et de ses abords et être l'occasion d'exprimer des signes de courtoisie qui potentiellement peuvent contribuer à faire évoluer une relation de la « connaissance » (*meeting*) vers l'amitié (*mating*) (chapitre 2). Après un examen des données relatives aux détections des co-présences, dont les statistiques descriptives sont présentées en Annexe (tableau A.5, p.164), nous privilégions un seuil correspondant à un volume horaire minimum de 12 heures cumulées de co-présences par vague d'observation, c'est-à-dire au minimum 1 heure par semaine.

Le graphique n°9 illustre les contacts de proximité entre les individus sous la forme de graphes non orientés. Sur ces graphes, chaque lien correspond à une co-présence une fois appliqué le seuil de 12 heures de co-présences par vague. Pour chaque vague, la distribution des degrés des nœuds est précisée, ainsi que la distribution par intervalles en heures pour la durée des contacts de proximité. L'algorithme utilisé pour le tracé des graphes *(layout algorithm*) est celui développé par Davidson et Harel (1996). Sur le graphique n°9, la taille des nœuds est proportionnelle à leurs degrés et l'épaisseur des liens est proportionnelle à la durée des échanges.

[15] Les graphes non seuillés sont présentés en Annexe (graphiques A.4, p.165).

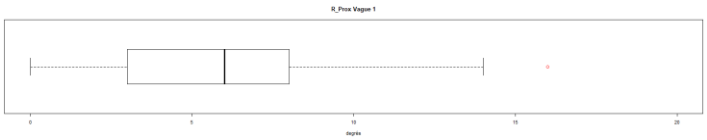
Graphique n°9 : représentation de l'évolution du réseau « R_Prox » (seuillé)



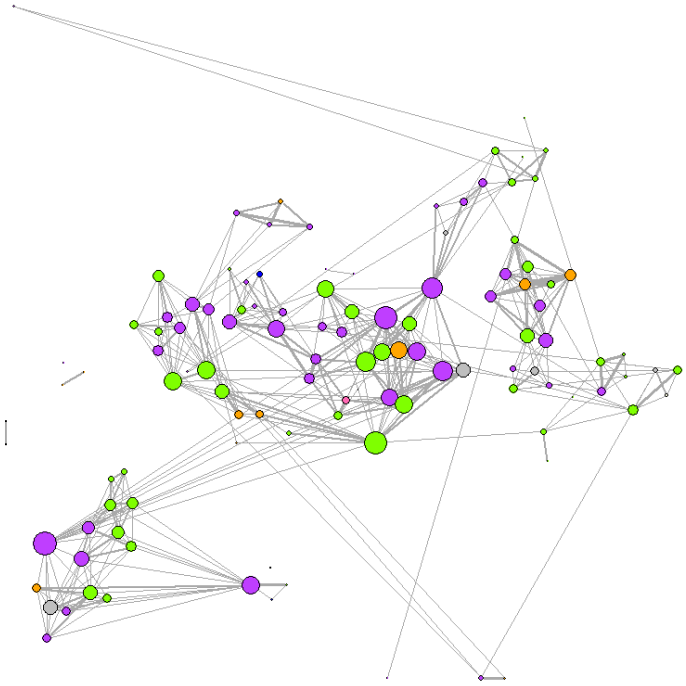
Vague n°1 « R_Prox»

(114 individus, 331 liens, seuil : ≥12h)

Distribution des degrés des nœuds



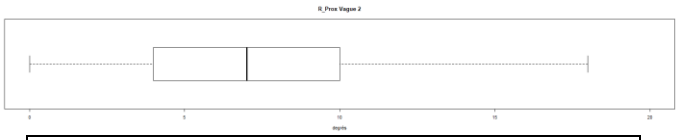
Distribution par intervalles de la durée des co-présences	
Intervalles (heures)	Effectifs (%)
[12 ;24)	23
[24 ;48)	28
[48 ;72)	24
[72 ;219]	25
moyenne : 54 h	100% (331 liens)



Vague n°2 « R_Prox»

(114 individus ; 404 liens, seuil : ≥12h)

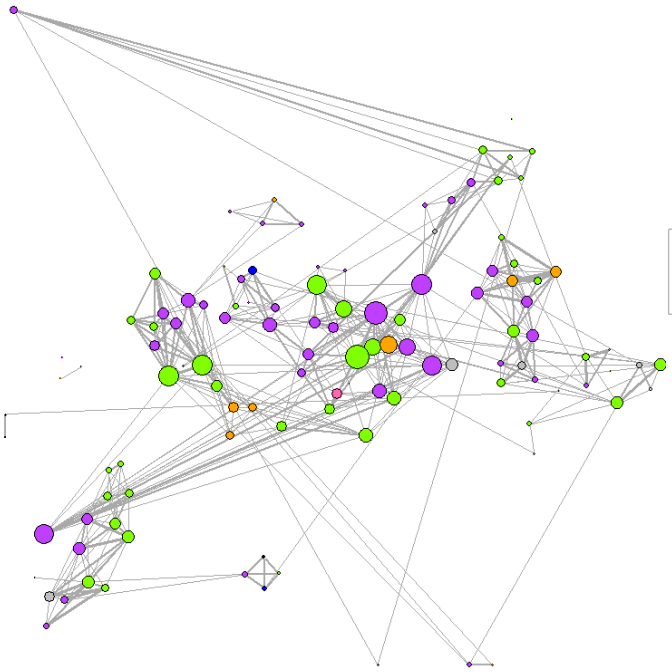
Distribution des degrés des nœuds



Distribution par intervalles de la durée des co-présences	
Intervalles (heures)	Effectifs (%)
[12;24)	19
[24;48)	23
[48;72)	15
[72;422]	43
moyenne : 83 h	100% (404 liens)

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

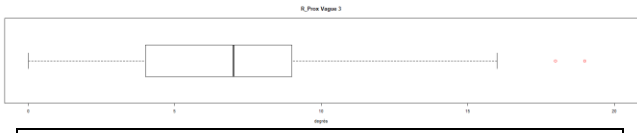
Graphique n°9 suite : représentation de l'évolution du réseau « R_Prox » (seuillé)



Vague n°3 « R_Prox»

(114 individus ; 398 liens, seuil : ≥12h)

Distribution des degrés des nœuds



Distribution par intervalles de la durée des co-présences	
Intervalles (heures)	Effectifs (%)
[12;24)	18
[24;48)	20
[48;72)	15
[72;380]	47
moyenne : 128 h	100% (398 liens)

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

Graphiques n°10 : distribution de la durée des co-présences du réseau « R_Prox». (intensité des liens : ≥ 12h ; effectif : 114 individus)

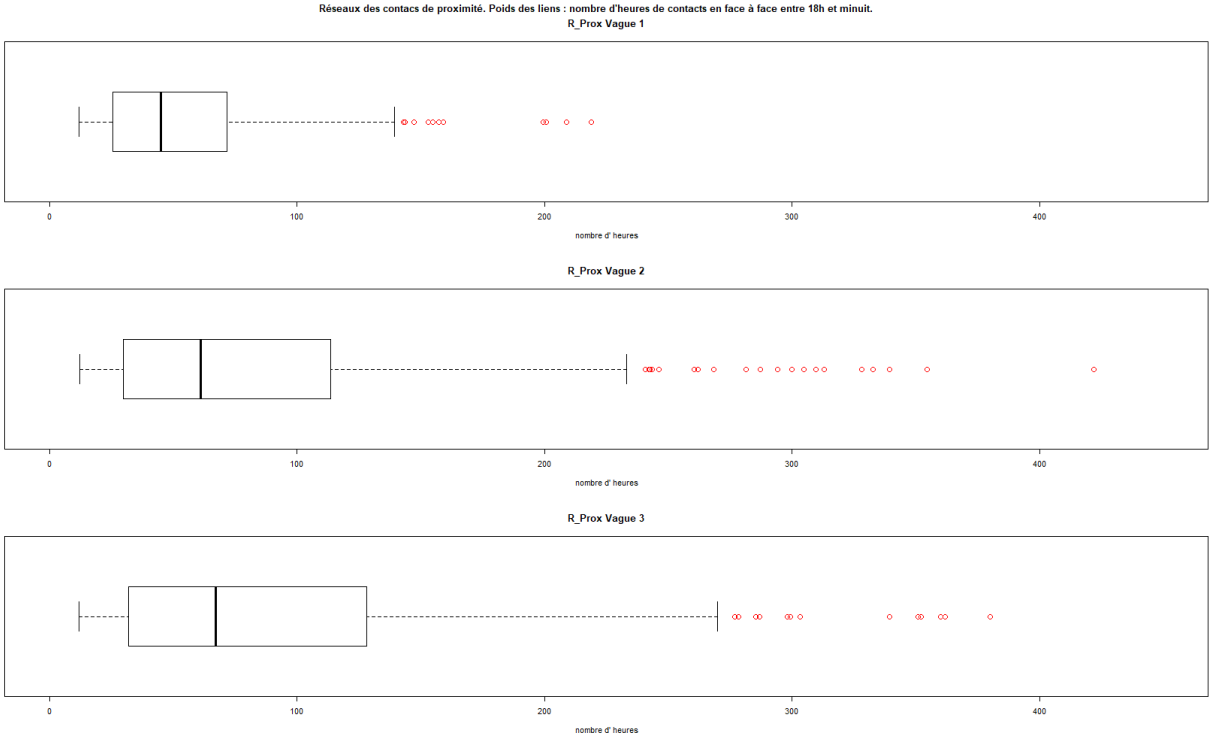


Tableau n°7 : métriques du réseau « R_Prox » seuillé, pour les 3 vagues d'observation

Réseau « R_Prox » non dirigé intensité des liens : co-présence ≥ 12h / vague	**Vague n°1**	**Vague n°2**	**Vague n°3**
Taille du réseau	114 nœuds, 331 liens	114 nœuds, 404 liens	114 nœuds, 398 liens
Densité	0,051	0,063	0,062
Longueur moyenne du chemin le plus court *(Average Path Length)*	3,565	4,117	5,222
Groupements connexes *(Clusters[16])* nombre [taille des groupes]	13 « clusters » [3, 15, 67, 4, 14, 2, 2, 1, 1, 1, 2, 1, 1]	8 « clusters » [2, 104, 2, 1, 1, 2, 1, 1]	7 « clusters » [2, 105, 1, 3, 1, 1, 1]
Communautés et modularité (*community structure detection based on edge betweenness)* (algorithme : « *cluster edge betweenness*[17] »)	19 « communautés » modularité : 0,76	14 « communautés » modularité : 0,72	13 « communautés » modularité :0,73
Coefficient global d'agglomération *Global Clustering Coefficient*	0,633	0,592	0,593
Coefficient local moyen d'agglomération *Local Clustering coefficient*	0,733	0,703	0,684
Centralisation de Freeman normalisée	0,090	0,096	0,106
Centralité de degré normalisée			
Moyenne	0,051	0,063	0,062
Ecart-type	0,031	0,038	0,036
Degré maximal	0,141	0,159	0,168
Centralité d'intermédiarité normalisée - valeur maximale *(standardized betweenness centrality*)	0,147	0,267	0,359
Origine ethnique des 10 premiers « *gatekeepers* »	Asiatique (6), Blanche (4)	Asiatique (4), Blanche (6)	Asiatique (5), Blanche (4), Métissée (1)

L'examen des graphes (graphique n°9) et des métriques (tableau n°7) du réseau « R_Prox » apportent un troisième éclairage descriptif au récit que nous tentons de reconstruire sur l'évolution entrelacée de la sociabilité amicale au sein de la communauté des étudiants au cours du temps.

Aucun des 3 graphes du réseau « R_Prox » n'est connexe et aucun des regroupements n'apparaît homogène en ce qui concerne l'origine ethnique, à l'exception de quelques binômes déconnectés de leurs pairs.

[16] Groupements détectés par l'algorithme « *Clusters* » (Csardi et Nepasz, 2005 ; boîte à outil « Igraph » du logiciel R).
[17] Algorithme élaboré par Csardi et Nepusz (2005) à partir des travaux de Newman et Girvan (2004).

Dans le premier graphe (vague n°1), visuellement nous distinguons grossièrement 3 principaux amas qui forment des groupements non connectés entre eux, dont un qui concentre une large majorité des individus. Au fil du temps, dès la deuxième vague ces 3 amas s'interconnectent et seuls quelques individus (entre 3 et 5) demeurent non reliés à l'un de ces amas. Confirmant cette tendance, de la première à la dernière vague le nombre de sous-graphes connexes maximum identifiés par l'algorithme « *clusters* » (Csardi et Nepusz, 2005) diminue de moitié et le principal groupement connexe s'étoffe de près de 40% d'individus entre ces 2 temps d'observation (tableau n°7). En outre, l'algorithme de détection de « communautés structurales » basé sur l'indice d'intermédiarité des liens (« *cluster edge betweenness* », Newman et Girvan, 2004 ; Csardi et Nepusz, 2005) met en évidence un nombre décroissant d'environ 1/3 de communautés structurales détectées au sein des graphes entre la première et la dernière observation. Ainsi, le nombre de modules très densément « intra-connectés » mais peu connectés aux autres modules tend à se réduire très nettement au cours du temps de vie de la communauté : **des liens, relatifs à la co-présence, de plus en plus nombreux se tissent entre des entre les individus de groupes initialement peu en interaction dans le même espace de proximité.**

Le coefficient d'agglomération globale chute de 4% entre la première et la dernière vague d'observation, avec une valeur moyenne d'environ 0,6 (tableau n°7). Il est un peu plus élevé que celui du réseau « R_Friend » (0,4 en moyenne) et 3 fois plus élevé que celui du réseau « R_Call » (0,17 en moyenne). L'évolution de la transitivité locale est également décroissante avec des valeurs moyennes les plus élevées de tous les réseaux analysées (0,7). Ce nombre plus élevé de triades transitives est très probablement expliqué par les co-présences physiques des individus lors d'activités partagées en groupes de taille plus ou moins importantes qui génèrent des détections multiples de présences, qui toutefois ne sont pas forcément des interactions dyadiques. Enfin, à l'examen de l'indice de centralité de Freeman (1979) le réseau « R_Prox » apparaît très peu centralisé. Il est constitué de multiples amas plus ou moins touffus ce qui contraste avec les réseaux « R_Call » et « R_Friend » pour lesquels cet indice est 3 à 4 fois plus élevé.

Pour le réseau « R_Prox », nous observons que les degrés maximums observés sont 3 à 6 fois plus faibles que ceux des 2 réseaux précédents avec, toutefois, une centralité de degré moyenne nettement plus élevée que celle du réseau « R_Call » (tableau n°7). Cependant, cette valeur est très en dessous du niveau moyen de centralité de degré du réseau des amitiés déclarées. **Ainsi, les individus les plus « actifs » et « populaires » au sein du réseau de proximité le sont beaucoup moins que ce qui apparaît au sein des réseaux des contacts téléphoniques et d'amitié.**

Ceci est probablement dû aux regroupements des individus lors d'activités où chacun est « détecté » co-présent avec un ou plusieurs pairs. Toutefois, en moyenne, les individus de cette communauté sont plus actifs dans leurs contacts en présentiel qu'à distance, ce qui semble cohérent compte-tenu de leur faible éloignement spatial. **En moyenne sur les 3 périodes d'observation du réseau « R_Prox », les individus ayant les plus hauts indices de centralité d'intermédiarité (normalisée) sont à égalité d'origine ethnique asiatique et blanche.**

Nous observons que les valeurs maximales de centralité d'intermédiarité sont inférieures à celles observées pour le réseau « R_Call » et celles mesurées pour les 2 premières vagues d'observation du réseau « R_Friend » (tableau n°7). Cependant, d'une part, lors de la dernière vague d'observation, les valeurs maximales de centralité d'intermédiarité du réseau des co-présences se rapprochent du niveau de celles observées pour les 2 autres réseaux. D'autre part, ce sont les seules valeurs maximales de centralité d'intermédiarité qui augmentent de façon continue, avec en outre une progression de plus de 60% entre la première et la dernière vague de mesure. **Les étudiants d'origines ethniques se répartissent à part presque égales près de la moitié des premiers rôles d'intermédiaires. En outre, avec le temps le rôle de ces « *gatekeepers* » progresse très sensiblement en importance au sein du réseau « R_Prox » et ce de façon inverse aux tendances observées pour les « *gatekeepers* » des réseaux « R_Call » et « R_Friend »**. L**ors de la dernière vague les valeurs du réseau des co-présences se rapprochent de celles observées pour les 2 autres réseaux.**

Synthèse

L'examen des métriques descriptives du réseau « R_Prox » révèle des différences singulières en comparaison avec les réseaux des relations d'amitiés déclarées et des contacts téléphoniques. Un élément important à considérer pour ce réseau est, d'une part, la distribution spatiale des individus au sein de la résidence et, d'autre part, leur participation en groupe aux activités proposées par la résidence familiale. Ainsi, il convient de prendre en compte qu'au sein du réseau « R_Prox » certaines détections de co-présences peuvent être le résultat d'évènements qui rassemblent de nombreux individus sans toutefois que cela implique un échange dyadique entre chacun d'entre eux, contrairement aux deux autres réseaux uniplexes qui sont basés sur des liens dyadiques effectifs (R_Call) ou déclarés (R_Friend).

4.2 Modélisation longitudinale des réseaux uniplexes et multiplexes

Les sections suivantes présentent les résultats des modélisations longitudinales pour :

- les 3 réseaux non-orientés uniplexes « R_Friend », « R_Call » et « R_Prox ».
- les 2 réseaux non-orientés multiplexes d'ordre 2 « R_Friend∩R_Call », « R_Friend∩R_Prox ».
- le réseau non-orienté multiplexe d'ordre 3 « R_Friend∩R_Call∩R_Prox ».

Au chapitre 3, nous avons détaillé d'une part, les spécifications des modèles et sa mise en œuvre l'algorithmique (SIENA) et, d'autre part, les choix des effets des variables socio-démographiques et structurales. Pour mémoire, les variables socio-démographiques sont les suivantes :

- le sexe, l'origine ethnique (A- Asiatique, B- Blanche) et le nombre d'enfants.

Les facteurs structuraux choisis pour les réseaux uniplexes et multiplexes sont les suivantes :

- la densité du réseau.
- l'appariement des individus de degré équivalent.
- l'activité et la popularité des individus.
- le nombre de triades transitives et le nombre de configurations du type « 4-cycles ».
- l'effet direct du réseau W sur le réseau X (*Network_X : Network_W*).
- la concordance des choix relationnels des acteurs du réseau X compte tenu de leurs choix relationnels au sein du réseau W (*agreement along W leading to X*).
- la tendance des individus à faire les mêmes choix relationnels au sein du réseau X que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau W (W leading *to agreement along X*).

Seuls les modèles qui satisfont aux 2 caractéristiques de bonne convergence de l'algorithme de simulation SIENA (« *t ratios* » < 0,1 et « ratio global maximum de convergence » < 0,25), sont présentés. La qualité d'ajustement des modèles (*goodness of fit*) n'est donnée que pour les meilleurs modèles obtenus à l'issue des simulations. Enfin, comme expliqué au chapitre 3, seules les simulations qui ont fourni un résultat en moins de 48h de calculs sont retenues dans cette étude.

4.2.1 Le réseau « R_Friend »

La première proposition de modélisation est réalisée à partir des données du réseau « R_Friend » avec un seuillage de l'intensité des liens à la valeur minimale de 3 (cf. chapitre 3). Les tableaux n°8 et n°9 ci-dessous présentent les indicateurs de base du réseau et les résultats de la modélisation pour le modèle du type « *Forcing & Initiative* ».

Tableau n°8 : évolution des indicateurs de base et de l'index de Jaccard pour le réseau « R_Friend » (intensité des liens : ≥ 3 ; 123 nœuds)

Vague d'observation	**n°1**	**n°2**	**n°3**	**n°4**
Densité	0,065	0,076	0,077	0,087
Degré moyen	7,967	9,285	9,333	10,618
Nombre de liens	490	571	574	653
Index de Jaccard	0,679	0,775	0,778	

Tableau n°9 : résultats des modélisations longitudinales pour le réseau « R_Friend » (123 nœuds)

Réseau « R_Friend » seuil : intensité des liens ≥ 3	**Modèles**					
	F1		**F2**		**F3**	
Effets	paramètre	(é.t)	paramètre	(é.t)	paramètre	(é.t)
taux 1	1,882	(0,139)	2,251	(0,188)	2,278	(0,193)
taux 2	1,289	(0,105)	1,428	(0,132)	1,441	(0,131)
taux 3	1,378	(0,116)	1,532	(0,136)	1,535	(0,134)
structuraux						
degré (densité)	–0,851***	(0,088)	–1,604***	(0,147)	–1,773***	(0,154)
appariement degré[1] équivalent			–0,152***	(0,029)	–0,138***	(0,028)
activité + popularité			0,064***	(0,012)	0,063***	(0,011)
triades transitives			0,435***	(0,028)	0,530***	(0,042)
4-cycles					–0,009**	(0,003)
co-variables						
sexe Alter	–0,093	(0,131)	0,014	(0,123)	–0,012	(0,125)
même sexe	–0,111	(0,087)	–0,109	(0,089)	–0,108	(0,084)
ethnie A Alter	1,120***	(0,215)	0,531*	(0,214)	0,547**	(0,208)
même ethnie A	–0,045	(0,108)	–0,121	(0,106)	–0,119	(0,107)
ethnie B Alter	0,716***	(0,210)	0,422*	(0,206)	0,424*	(0,194)
même ethnie B	–0,292**	(0,106)	–0,268*	(0,106)	–0,270**	(0,104)
nombre d'enfants d'Alter	0,213*	(0,100)	0,366***	(0,096)	0,342***	(0,099)
similarité du nombre d'enfants	–0,018	(0,139)	–0,011	(0,137)	–0,005	(0,137)
inter. sexe * nbre d'enfants	–0,086	(0,173)	–0,195	(0,164)	–0,197	(0,159)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001; (é.t) : écart-type.
(1) racine carrée des degrés.

Examen des métriques descriptives

Les éléments du tableau n°8 montrent que les degrés moyens du réseau progressent à chaque vague et que les index de Jaccard sont nettement au-dessus de la valeur de référence 0,3 : le réseau évolue, avec des créations et des dissolutions de liens, mais sans excès. La densité moyenne du réseau (0,07) est nettement plus faible que celle observée pour le réseau non seuillé (0,16), ce qui est dû à l'application d'un seuil de 3 pour l'intensité des liens. Ce seuil, qui permet de filtrer les relations de faible intensité d'amitié, a mécaniquement éliminé près de la moitié des liens qui relèvent plus de la connaissance que de l'amitié, ce qui a pour conséquence de réduire la densité du réseau initial.

Examen des résultats des modélisations

La première simulation présentée au tableau n°9 (F1) est réalisée sans effets structuraux, à l'exception de celui de densité systématiquement inclus et non interprété dans les modèles (cf. chapitre 3). Ce premier modèle, que nous nommons « naïf », comprend donc uniquement les effets des co-variables socio-démographiques et ne prend pas en compte l'impact des facteurs structuraux.

Le modèle naïf (F1)

Plusieurs variables relatives aux origines ethniques des individus avec qui Ego déclare nouer une relation amicale ont des effets statistiquement très significatifs. L'origine ethnique d'Alter apparaît jouer un rôle important dans l'évolution temporelle de ces nominations.

Les effets positifs relatifs à l'« ethnie A » d'Alter et à l'« ethnie B » d'Alter montrent que la popularité relative à l'appartenance aux 2 ethnies considérées est significativement forte dans l'évolution des relations amicales. Toutefois, en ce qui concerne l'effet d'identité dyadique qui traduit l'homophilie, seul l'effet - négatif - relatif à la même appartenance à « l'origine ethnique B » est très significatif. Les individus qui partagent cette origine sont beaucoup moins susceptibles, au fil du temps, de déclarer des liens d'amitié entre eux que de faire des choix alternatifs.

Enfin, la dernière variable socio-démographique qui a un effet significatif est celle relative au « nombre d'enfants » d'Alter. L'effet est significativement positif : le fait d'avoir des enfants contribue positivement à l'évolution des nominations. Il s'agit d'une variable de contrôle dont le résultat est attendu : la présence d'enfants joue un rôle non négligeable dans le développement des relations entre les foyers. En effet, le comportement relationnel n'est pas le même pour des individus avec ou sans enfants : les contraintes horaires sont différentes et les opportunités d'activités conjointes avec d'autres familles et donc d'établir des relations sont plus nombreuses. Les effets des autres variables, dont l'homophilie relative au sexe, n'apparaissent pas statistiquement significatifs aux seuils considérés.

Introduction des facteurs structuraux (F2, F3)

Les variables socio-démographiques

Pour les modèles F2 et F3, l'introduction des facteurs structuraux réduit sensiblement la significativité des effets liés à l'origine ethnique des individus. La popularité des 2 ethnies perd en force explicative de l'évolution des citations de relations d'amitié. De surcroît, l'homophilie relative à l'« ethnie B » perd également du pouvoir explicatif.

Les signes des effets significatifs relatifs aux origines ethniques ne changent pas mais les valeurs respectives[18] de leurs coefficients évoluent différemment du modèle F1 aux modèles F2 et F3. En effet, l'amplitude des coefficients des effets relatifs à la popularité de citation de l'« ethnie A » et de l'« ethnie B » dans les relations d'amitiés diminue respectivement de moitié. Par contraste, l'amplitude du coefficient de l'effet d'identité dyadique relatif à l'homophilie d'« ethnie B » change très peu.

En ce qui concerne la co-variable « nombre d'enfants » d'Alter, d'une part, sa significativité statistique augmente très nettement et, d'autre part, la valeur de son coefficient croît sensiblement du modèle F1 aux modèles F2 et F3 : la popularité liée au nombre d'enfants augmente en significativité positive dans la nomination des relations d'amitiés. L'introduction des effets structuraux a tendance à révéler l'importance de ce facteur plus que de capter son effet.

Les effets structuraux

Tous les effets structuraux introduits sont significatifs au seuil de significativité le plus exigeant. L'effet relatif à la transitivité relationnelle sous la forme de triades est très significatif et positif. La dynamique « les amis de mes amis sont mes amis » est très prégnante dans l'évolution des nominations des relations d'amitiés, ce qui est un phénomène attendu dans ce type de réseau d'amitié. En revanche, la configuration structurale transitive plus complexe de la forme « 4-cycles » est très significativement négative : on peut souligner un déséquilibre transitif dans cette configuration qui implique 4 individus dans des relations croisées. La transitivité relationnelle apparaît ainsi plus complexe que celle exprimée par le seul décompte du nombre de triades transitives.

[18] La comparaison des paramètres des différents effets est possible étant donné la forme linéaire de la fonction objective. Toutefois, il faut prendre en compte que les coefficients présentés ne sont pas standardisés, ce qui nécessite de développer un calcul avec les coefficients standardisés des différents effets pour pouvoir les comparer. Néanmoins, la comparaison relative des coefficients d'un même effet, d'un modèle à l'autre, ne nécessite pas cette opération.

Enfin, les deux derniers effets structuraux gagnent à être interprétés ensemble : l'effet d'activité/popularité structurale et celui relatif à la préférence des individus à se lier avec des pairs de degré équivalent. Ces 2 effets sont très significatifs mais de signes opposés. L'effet très significativement négatif d'appariement de degré équivalent nous conduit à déceler une évolution vers l'hétérophilie du point de vue de la connectivité des individus dans le réseau des relations amicales. Ainsi, on observe une recherche de pairs peu connectés par des individus très connectés et inversement. L'effet de popularité/activité structurale, positif, bénéficie de cette évolution : les individus actifs et populaires dans ce réseau de nomination sont plus susceptibles d'être nommés comme amis selon une dynamique de renforcement.

Synthèse

La seule prise en compte de facteurs exogènes au réseau nous aurait conduit à sur-interpréter l'importance du phénomène d'homophilie expliqué par les attributs socio-démographiques des individus. Ce phénomène demeure notable en ce qui concerne les origines ethniques des individus mais il perd en significativité lors de l'introduction de facteurs structuraux endogènes au réseau des nominations d'amis.

Parmi les facteurs structuraux introduits, celui relatif à l'appariement des individus de degrés équivalent montre une évolution du réseau plus expliquée par la recherche de pairs d'un niveau de connectivité dissemblable que d'attributs socio-démographiques similaires ou dissimilaires. L'évolution du réseau n'apparaît donc pas, au regard des facteurs endogènes, structurée par l'homophile. On observe plutôt une recherche d'hétérogénéité structurale, plus précisément un intérêt des acteurs à se lier avec des individus qui occupent une variété de positions structurales, exprimées par les degrés de connectivité dissimilaires.

On observe donc, dans l'évolution des liens d'amitiés déclarés, l'existence persistante d'un phénomène d'homophilie relative à l'origine ethnique mais celle-ci est atténuée par la prise en compte des jeux relationnels traduits par des facteurs endogènes aux réseaux, dont notamment la transitivité relationnelle et, de façon plus intéressante dans le cadre de nos hypothèses, un déplacement des relations vers plus d'hétérogénéité structurale du point de vue de la connectivité.

Qualité d'ajustement des modèles

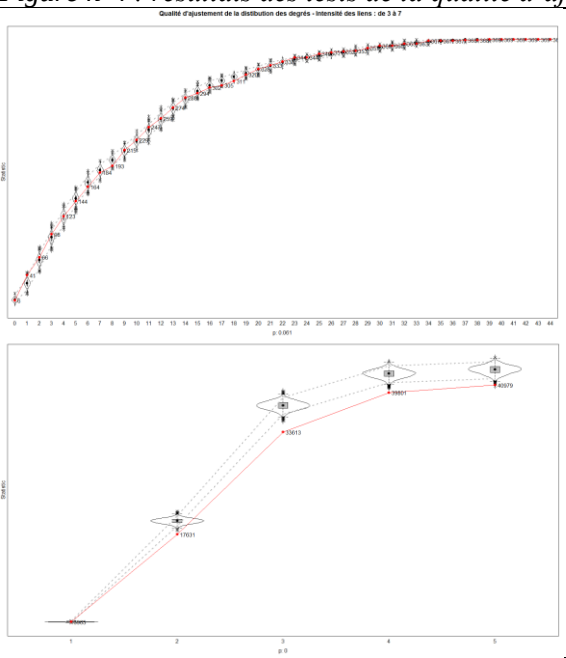
La qualité d'ajustement des modèles simulés avec un seuil d'intensité des liens à 3 est appréciée par la distribution des distances géodésiques et des degrés (cf. chapitre 3). La figure n°4 présente les résultats de différents tests d'appréciation de la qualité du modèle C3 formalisé comme ci-dessous :

R_Friend (seuil d'intensité des liens ≥ 3) ~ Degrés +Appariement par degré
+ Σ Popularité & Activité + Triades transitives + Cycle-4
+ co-variables d'Alter (sexe, ethnie, nbre d'enfants)
+ égalité (sexe, ethnie) + similarité (nbre d'enfants) + interaction sexe * nbre d'enfants.

Ce modèle fournit les meilleurs résultats d'ajustement parmi les modèles simulés. Les degrés du modèle empirique (points rouges) se situent pour la plupart dans les limites de la distribution des degrés. En revanche, ce n'est pas le cas pour les distances géodésiques de 2 à 5 (la valeur 1 correspond aux liens directs). En effet, le nombre de dyades simulées avec de petites distances apparaît trop élevé : le réseau simulé est « trop connecté » en comparaison avec le réseau empirique. Pour cet indicateur, l'ajustement est par conséquent relativement médiocre.

Le test d'hétérogénéité temporelle permet de vérifier que ces résultats ne sont pas la conséquence d'une hétérogénéité temporelle. Sur la représentation graphique du test, les bandes rouges correspondent aux intervalles de confiance des paramètres du modèle pour les vagues n°2 et n°3, et les bandes grises aux intervalles de confiance pour la vague n°1. Pour la plupart, les bandes rouges se superposent aux bandes grises. De surcroît, les statistiques de référence (densité et distribution des degrés) sont assez proches pour toutes les vagues (cf. section 4.2). Nous concluons que les paramètres ne varient pas systématiquement d'une période à l'autre et nous pouvons ainsi estimer un seul modèle pour les 4 vagues du réseau « R_Friend ».

Figure n°4 : résultats des tests de la qualité d'ajustement du modèle F3

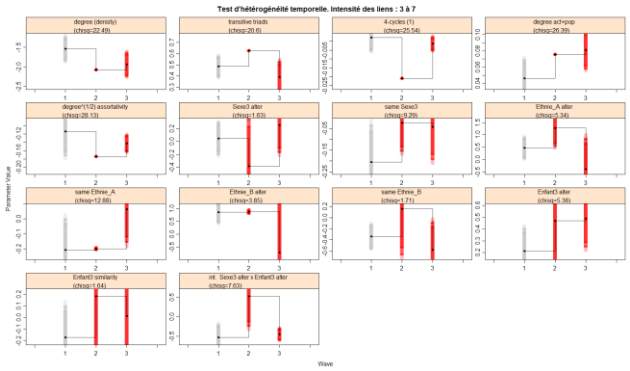


Degrés.

p-value = 0,061

Distance.

p-value = 0



Test d'hétérogénéité temporelle.

4.2.2 Le réseau « R_Call »

La deuxième proposition de modélisation est réalisée à partir des données du réseau « R_Call », avec un seuillage de l'intensité des liens à la valeur de 3 appels par vague au minimum (cf. chapitre 3). Les tableaux n°10 et n°11 ci-dessous présentent les indicateurs de base du réseau et des résultats de la modélisation pour le modèle du type « *Forcing & Initiative* ».

Tableau n°10 : évolution des indicateurs de base pour le réseau « R_Call »
(intensité des liens : ≥ 3 appels/vague; 112 nœuds)

Vague d'observation	**n°1**	**n°2**	**n°3**
Densité	0,030	0,032	0,029
Degrés moyens	3,393	3,571	3,268
Nombre de liens	190	200	183
Index de Jaccard	0,681	0,672	

Tableau n°11 : résultats des modélisations longitudinales pour le réseau « R_Call » (112 nœuds)

Réseau « R_Call » seuil : intensité des liens ≥ 3 appels/vague	**Modèles**					
	C1		**C2**		**C3**	
Effets	paramètre	(é.t)	paramètre	(é.t)	paramètre	(é.t)
taux 1	0,768	(0,097)	0,834	(0,108)	0,828	(0,108)
taux 2	0,780	(0,096)	0,835	(0,110)	0,830	(0,104)
structuraux						
degré (densité)	–1,959***	(0,182)	–2,096***	(0,302)	–1,964***	(0,325)
appariement degré[1] équivalent			–0,191*	(0,087)	–0,234*	(0,094)
activité+popularité			0,050*	(0,022)	0,058*	(0,024)
triades transitives			1,120***	(0,154)	0,960***	(0,199)
4-cycles					0,058	(0,056)
co-variables						
sexe Alter	0,097	(0,293)	0,286	(0,334)	0,270	(0,341)
même sexe	–0,295†	(0,178)	–0,382*	(0,193)	–0,370†	(0,189)
ethnie A Alter	0,401	(0,332)	0,701	(0,431)	0,686	(0,440)
même ethnie A	0,079	(0,195)	0,118	(0,207)	0,108	(0,199)
ethnie B Alter	0,111	(0,339)	0,558	(0,429)	0,542	(0,419)
même ethnie B	0,178	(0,190)	0,167	(0,206)	0,179	(0,200)
nombre d'enfants d'Alter	0,381*	(0,182)	0,420*	(0,185)	0,457*	(0,184)
similarité du nombre d'enfants	0,068	(0,267)	0,127	(0,261)	0,140	(0,270)
inter. sexe * nbre d'enfants	–0,182	(0,326)	–0,353	(0,375)	–0,333	(0,368)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001; (é.t) : écart-type.
(1) racine carrée des degrés.

Examen des métriques descriptives

A chaque vague d'observation du réseau des contacts téléphoniques, les degrés moyens du réseau progressent et les index de Jaccard sont nettement au-dessus de la valeur de référence 0,3 : le réseau évolue, sans excès de création ou de dissolution des liens (tableau n°10). La densité moyenne du réseau (0,03) est nettement plus faible que celle observée pour le réseau « R_Friend » commenté précédemment (0,076). Les liens construits à partir des contacts téléphoniques forment un réseau plus distendu que le réseau des nominations d'amis. On note ici une probable conséquence d'un effet relatif à la structure d'opportunité : les individus vivent dans une même résidence et étudient sur le même campus, certains dans un même département d'études. Par conséquent, l'échange entre amis par téléphone est une modalité d'interaction moins fréquente que dans le cas d'un éloignement géographique important. Dans cette situation, d'autres modalités de contacts plus immédiats sont privilégiées, en particulier celle relative aux co-présences dont nous verrons par la suite la consistance.

Le modèle naïf (C1)

La première simulation (tableau n°11, modèle C1) comprend uniquement les effets des co-variables socio-démographiques. Dans ce modèle naïf, seuls les effets relatifs au « nombre d'enfants » d'Alter et à l'identité dyadique liée au sexe sont statistiquement significatifs et de signes différents. L'effet dû à l'homophilie de sexe est significativement négatif à un niveau de significativité faible : ce phénomène d'homophilie joue un faible rôle négatif dans la tendance des individus du même sexe à interagir entre eux par téléphone mais l'effet est modeste. L'effet significativement positif de la popularité du nombre d'enfants dans les échanges par téléphone relève probablement d'intérêts communs liés aux soins, à l'éducation et à l'organisation d'activités familiales partagées pour les enfants. Cet effet est attendu et ce facteur joue le rôle de variable de contrôle dans le modèle.

Introduction des facteurs structuraux (C2, C3)

Les variables socio-démographiques

Nous observons que l'introduction des facteurs structuraux modifie relativement peu l'impact significatif des co-variables socio-démographiques observées dans le modèle C1. Nous soulignons la persistance de l'hétérophilie relative au sexe, qui par ailleurs augmente en significativité dans le modèle C2 lors de l'introduction des facteurs structuraux. Les contacts par téléphone traduisent une sociabilité différente au regard du sexe que celle exprimée par le réseau des amitiés déclarées pour lequel cette variable n'était pas significative.

Les effets structuraux

Tous les effets structuraux, à l'exception du facteur de transitivité relationnelle lié au dénombrement de formes « 4-cycles », sont statistiquement significatifs mais à des seuils différents.

Parmi ceux-ci, le facteur qui traduit les relations transitives sous la forme de triades a un effet très positivement significatif sur l'évolution du réseau des interactions à distance. Ici encore, la dynamique « j'interagis à distance avec les amis de mes amis » est un élément fort dans l'évolution du réseau.

Dans une moindre mesure, l'effet d'appariement de degré équivalent et celui relatif à l'activité et à la popularité sont également statistiquement significatifs. Les différences de signes sont les mêmes que celles observées pour le réseau des relations amicales : d'une part, on note de nouveau une évolution vers l'hétérophilie du point de vue de la connectivité dans les relations à distance entre les individus et, d'autre part, l'effet combiné de popularité et d'activité contribue positivement à la dynamique des interactions distancielles.

Synthèse

Dans ce réseau seule une hétérophilie relative au sexe est observée. Son effet significatif est modeste dans le modèle « naïf » mais son pouvoir explicatif progresse un peu lors de l'inclusion des facteurs structuraux dans le modèle.

Plus intéressant dans le cadre de nos hypothèses, nous observons dans ce réseau comme dans le précédent un déplacement des relations vers plus d'hétérogénéité dans l'appariement des individus au regard de leur connectivité. L'effet est nettement plus modéré dans le réseau « R_Call » que dans le réseau « R_Friend », mais il apporte un élément qui confirme la tendance d'une évolution du réseau expliquée par la recherche de pairs dissemblables en termes structuraux. A l'examen des facteurs endogènes, l'évolution du réseau des relations effectives, observées par les contacts à distance, n'apparaît pas ici non plus fortement structurée par une tendance à l'homophilie.

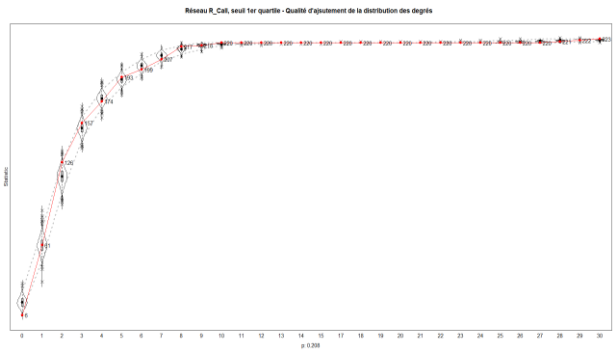
Qualité d'ajustement des modèles

La qualité d'ajustement des modèles simulés, avec un seuillage de l'intensité à 3 appels par vague, est évaluée par la distribution des distances géodésiques et des degrés (cf. chapitre 3). La figure n°5 présente les résultats du modèle C2 tel que formalisé ci-dessous :

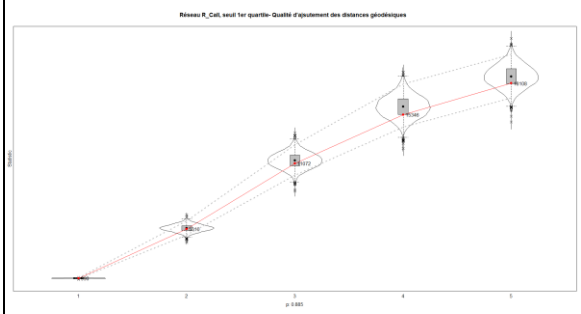
R_Call (seuil intensité des liens ≥ 1er q.) ~ Degrés +Appariement par degré
+ Σ Popularité & Activité + Triades Transitives
+ co-variables d'Alter (sexe, ethnie, nbre d'enfant)
+ égalité (sexe, ethnie) + similarité (nbre d'enfants) + interaction sexe * nbre d'enfants.

Ce modèle fournit les meilleurs résultats d'ajustement parmi les modèles testés. Les degrés et les distances géodésiques du modèle empirique se situent très correctement dans les limites de la distribution des degrés et des distances géodésiques du modèle simulé. Les deux « *p-value* » confirment la bonne qualité d'ajustement. Le test d'hétérogénéité temporelle montre que ces résultats ne sont pas la conséquence d'une hétérogénéité temporelle. Nous concluons que les paramètres ne varient pas systématiquement d'une période à l'autre et qu'un seul modèle permet de simuler l'évolution du réseau pour les 3 vagues.

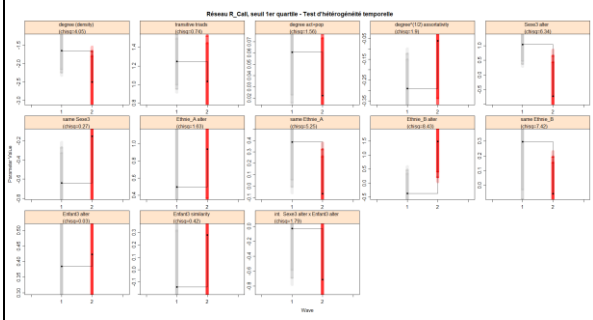
Figure n°5 : résultats des tests de la qualité d'ajustement du modèle C2



Degrés. p-value = 0,208



Distance. p-value = 0,885



Test d'hétérogénéité temporelle

4.2.3 Le réseau « R_Prox »

La troisième proposition de modélisation est réalisée à partir des données du réseau « R_Prox ». Un seuil de 12 heures de co-présence par vague au minimum est appliqué sur l'intensité des liens (cf. chapitre 3). Les tableaux n°12 et n°13 ci-dessous présentent les indicateurs de base du réseau et les résultats de la modélisation pour modèle du type « *Forcing & Initiative* ».

Tableau n°12 : évolution des indicateurs de base pour le réseau « R_Prox »
(seuil : intensité des liens : ≥ 12 h de co-présence/vague; 114 nœuds)

Vague d'observation	n°1	n°2	n°3
Densité	0,051	0,063	0,062
Degré moyen	5,807	7,088	6,982
Nombre de liens	331	404	398
Index de Jaccard	0,713	0,774	

Tableau n°13 : résultats des modélisations longitudinales pour le réseau « R_Prox » (114 nœuds)

Réseau « R_Prox » seuil : ≥12h de co-présence/vague	Modèles					
	P1		P2		P3	
Effets	paramètre	(é.t)	paramètre	(é.t)	paramètre	(é.t)
taux 1	1,213	(0,113)	1,793	(0,211)	1,780	(0,220)
taux 2	0,983	(0,101)	1,297	(0,151)	1,303	(0,149)
structuraux						
degré (densité)	–1,234***	(0,128)	–1,146***	(0,248)	–1,238***	(0,235)
appariement degré[1] équivalent			–0,162***	(0,026)		
activité+popularité					–0,071***	(0,011)
triades transitives			0,838***	(0,049)	0,811***	(0,047)
co-variables						
sexe Alter	–0,246	(0,183)	–0,251	(0,198)	–0,264	(0,192)
même sexe	0,066	(0,134)	0,096	(0,132)	0,094	(0,132)
ethnie A Alter	0,603*	(0,302)	0,585†	(0,328)	0,580†	(0,299)
même ethnie A	0,147	(0,159)	0,165	(0,158)	0,168	(0,154)
ethnie B Alter	0,546†	(0,315)	0,449	(0,300)	0,431	(0,296)
même ethnie B	–0,106	(0,172)	–0,102	(0,167)	–0,105	(0,159)
nombre d'enfants d'Alter	0,042	(0,149)	0,169	(0,152)	0,160	(0,140)
similarité du nombre d'enfants	0,070	(0,204)	0,173	(0,206)	0,164	(0,198)
inter. sexe * nbre d'enfants	–0,112	(0,285)	0,138	(0,275)	0,101	(0,259)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001; (é,t) : écart-type.
(1) racine carrée des degrés. Note : l'introduction simultanée du facteur nombre de triades transitives et des facteurs relatifs à l'appariement de degré équivalent et à l'activité + popularité conduisent à une mauvaise convergence de l'algorithme. Ils sont donc introduits différentiellement (P2, P3). Pour le facteur « 4-cycles », l'algorithme ne donnant pas de résultats dans les 48h il est exclu.

Examen des métriques descriptives

L'examen des degrés moyens du réseau progressent et les index de Jaccard nous permettent d'apprécier que le réseau évolue sans fortes variations (tableau n°12). La densité moyenne du réseau (0,058) est proche de celle du réseau « R_Friend » (0,076) et 2 fois plus élevée que celle observée pour le réseau « R_Call » (0,03). Elle est nettement plus faible que la densité observée pour le réseau non seuillé (0,486), ce qui est dû à l'application d'un seuil de 12 heures de co-présence pour l'intensité des liens par vague d'observation (soit au minimum 1h / semaine). Ce seuil permet de filtrer les co-présences de faibles durées que nous considérons dans le cadre de notre étude comme peu significatifs en termes d'échange relationnel du registre de l'amitié. La conséquence est une réduction d'un facteur de 10 du nombre de liens et de la densité du réseau initial. Un choix différent de seuil serait très certainement fait dans le cadre d'une étude épidémiologique.

Le modèle naïf (P1)

Le premier modèle (P1, tableau n°13), comprend uniquement les effets des co-variables socio-démographiques. Les effets de la variable « ethnie A » d'Alter et, à un seuil de significativité moins élevé, de « ethnie B » d'Alter sont statistiquement significatifs et leurs coefficients sont positifs. Ces effets, liés à la popularité de l'origine ethnique des individus avec qui Ego a des échanges de proximité, jouent un rôle non négligeable dans l'évolution du réseau. L'effet est plus prononcé pour l'origine ethnique asiatique.

Les autres variables socio-démographiques n'ont pas d'effets statistiquement significatifs aux seuils de significativité fixés dans cette étude.

Introduction des facteurs structuraux (P2, P3)

Les variables socio-démographiques

Pour le réseau « R_Prox », l'introduction des facteurs structuraux (modèles P2 et P3, tableau n°13) réduit la significativité des effets relatifs à l'« ethnie A » d'Alter, mais le coefficient du paramètre reste sensiblement de même ampleur que celle observée au sein du modèle naïf (P1). En revanche, l'effet de la variable « ethnie B », faiblement significatif dans le modèle P1, n'est plus significatif. Cela tend à souligner la prégnance d'un effet lié à l'origine ethnique asiatique qui, bien qu'affaibli dans son pouvoir explicatif de l'évolution du réseau, ne disparait pas dans le modèle qui comprend des facteurs endogènes structuraux.

Les autres variables socio-démographiques n'ont pas d'effets statistiquement significatifs.

Les effets structuraux

Les facteurs structuraux introduits différentiellement dans les modèles P2 et P3 ont tous des effets statistiquement très significatifs : ceux relatifs au nombre de triades transitives, à l'appariement de degré équivalent et à l'activité combinée à la popularité structurale. Pour des problèmes de convergence de l'algorithme de simulation, ces deux derniers facteurs ne sont pas associés ensemble avec le facteur relatif au nombre de triades transitives. Les effets liés à l'appariement de degré équivalent et à l'activité combinée à la popularité ont, dans leur modèle respectif, des coefficients de signes négatifs. De nouveau, on peut proposer comme lecture de l'effet d'appariement une évolution vers de l'hétérophilie du point de vue de la connectivité des individus.

En revanche, pour l'effet combiné de l'activité et de la popularité, il semble que, selon les résultats du modèle P3 où cet effet est dissocié de celui de l'appariement, l'interprétation n'est plus la même que pour les 2 réseaux précédents : la dynamique liée à cet effet apparaît ici inverse. La combinaison de l'activité et de la popularité structurales joue ici un rôle négatif dans l'explication de l'évolution des relations de proximité : une recherche de contacts en co-présences avec des individus peu structuralement actifs et populaires apparaît émerger, avec une tendance à l'évolution vers l'hétérophilie sur ces dimensions.

Synthèse

La modélisation basée sur la détection des co-présences des individus souligne une seule variable socio-démographique - l'origine ethnique asiatique - dont l'effet est persistant dans le pouvoir explicatif de l'évolution du réseau. Toutefois, une fois les facteurs structuraux endogènes au réseau introduits cet effet devient très faiblement explicatif.

En revanche, tous les effets structuraux sont très significatifs aux seuils de significativité les plus élevés. Le phénomène d'hétérogénéité expliqué par le facteur d'appariement de degré dissimilaires, déjà observé dans les simulations des 2 autres réseaux, se confirme, mais la détection des co-présences semble révéler une recherche de contacts dissimilaires du point de vue de l'activité et la popularité structurales.

Le modèle qui restitue les co-présences est certainement le plus compliqué à interpréter. En effet, comme nous l'avons souligné lors des analyses descriptives (chapitre 3), l'information restituée par les détecteurs de proximité des individus n'impliquent pas forcément une relation dyadique. Les individus sont identifiés comme étant spatialement proches, mais l'établissement d'une interaction dyadique entre les individus n'est pas systématique.

La détection des co-présences peut être en partie révélatrice de l'allocation des appartements aux étudiants. En effet, cela peut privilégier le regroupement d'étudiants originaires du continent asiatique par exemple. Par conséquent, ceci ne relèverait pas d'un choix de relation mais d'un effet induit par la structure qui contraindrait les opportunités de contacts (cf. chapitre 2). La question serait alors de savoir si les individus se comporteraient différemment s'ils étaient placés dans un autre contexte, une structure d'opportunité différente. Toutefois, les informations fournies par le MIT sur la résidence familiale mettent en valeur les nombreuses activités proposées aux résidents dans l'objectif explicite de favoriser les interactions afin de créer une communauté vivante et dynamique[19]. Ainsi, des structures d'opportunités organisées sont susceptibles de favoriser le décloisonnement résidentiel, ce qui pourrait modérer l'effet relatif à la proximité des hébergements. Néanmoins, il n'apparait pas possible en l'état, avec les données disponibles, de discerner ce qui peut relever d'une part, d'un choix de co-présence subi du fait des contraintes de l'organisation spatiale et, d'autre part, d'un choix de contact dyadique.

Les 3 réseaux uniplexes analysés sont de nature différente dans la restitution des relations sociales entre les individus suivis :

- le réseau « R_Friend » restitue des relations d'amitiés dyadiques déclarées.
- le réseau « R_Call » restitue des contacts dyadiques factuels à distance.
- le réseau « R_Prox » restitue des co-présences factuelles, pas forcément dyadiques.

Ils expriment sans surprise de façon contrastée les relations sociales entre les individus. Nous observons que ces 3 réseaux ne sont pas redondants dans leur récit de l'évolution des relations sociales entretenues par les individus au cours des 9 mois d'observation. Ils sont, bien au contraire, complémentaires : ils permettent de croiser les relations d'amitié déclarées avec des interactions distancielles effectives et des co-présences. L'étude des influences mutuelles de ces croisements par la modélisation des réseaux multiplexes est l'objet de la section suivante.

[19] Le lecteur pourra constater la grande variété des activités proposées aux étudiants du MIT de la résidence « *Wesgate familly housing* » sur le site dédié : http://westgate.mit.edu/.

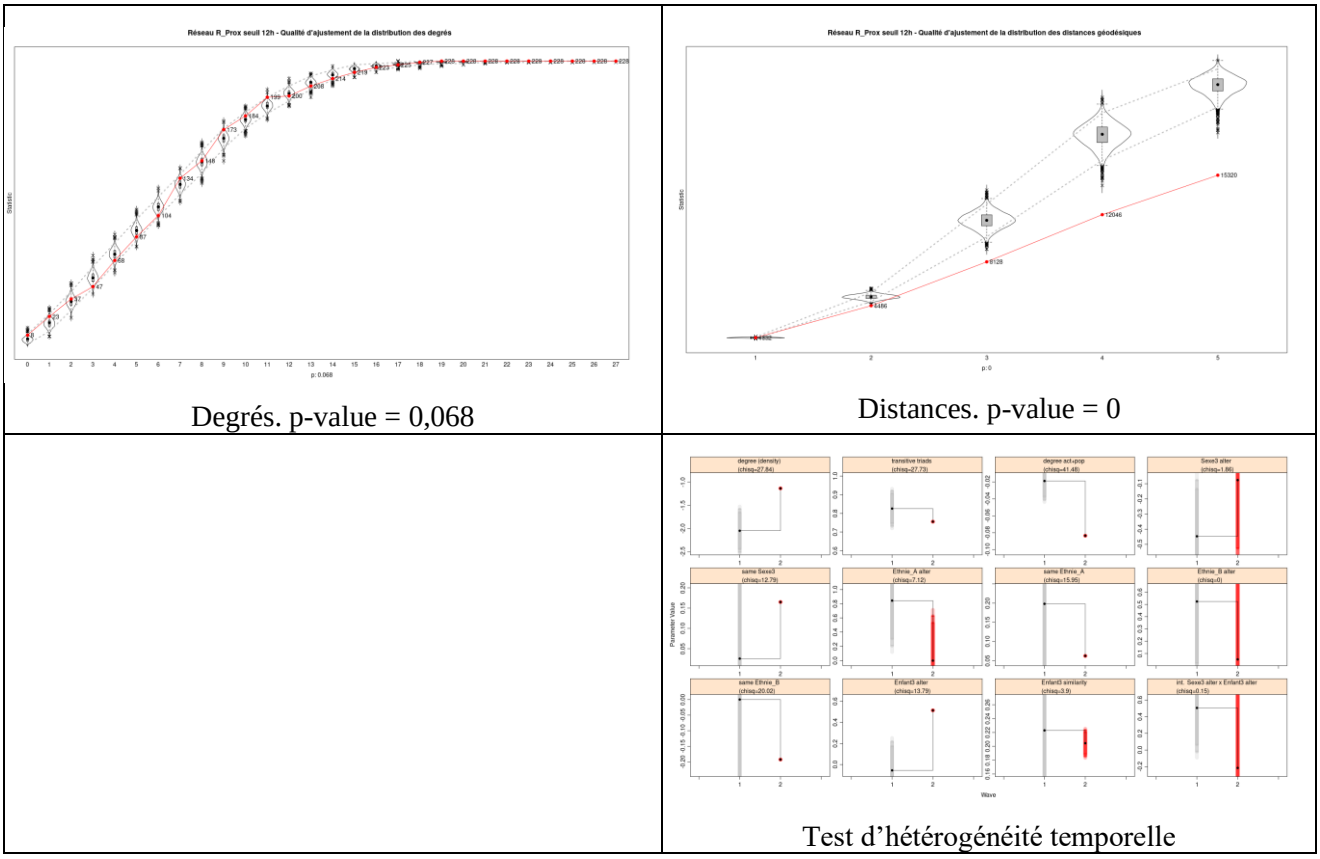
Qualité d'ajustement des modèles

La qualité d'ajustement des modèles simulés avec un seuillage de l'intensité est évaluée par la distribution des distances géodésiques et des degrés (cf. section 4.2). La figure présente les résultats de la qualité d'ajustement pour le modèles P2 ci-dessous :

R_Prox (seuil : ≥ 12h de co-présence /vague) ~ Degrés + Appariement de degré + Triades transitives
+ co-variables d'Alter (sexe, ethnie, nbre d'enfant)
+ égalité (sexe, ethnie) + similarité (nbre d'enfants) + interaction sexe * nbre d'enfants.

Les distributions des degrés pour les modèles simulés P2 et P3 sont très correctes au regard des valeurs observée pour les modèles empiriques. En revanche, les distributions des distances géodésiques des modèles simulés sont systématiquement plus élevées que les valeurs observées. Les tests d'hétérogénéité temporelle montrent que, globalement, ces résultats ne sont pas la conséquence d'une hétérogénéité temporelle. Nous concluons que les paramètres ne varient pas systématiquement d'une période à l'autre et ces modélisations permettent de simuler l'ensemble de l'évolution du réseau.

Figure n°6 : résultats des tests de la qualité d'ajustement du modèle P2



Degrés. p-value = 0,068

Distances. p-value = 0

Test d'hétérogénéité temporelle

Les résultats de la qualité d'ajustement pour le modèle P3 sont en Annexe (figure A.7, p.172).

4.2.4 Le réseau « R_Friend ∩ R_Call »

La première modélisation multiplexe est réalisée à partir de la combinaison « R_Friend ∩ R_Call ». Les seuils appliqués sur l'intensité des liens sont les suivants (cf. chapitre 4.2.2.) :

- une valeur minimale de 3 pour l'intensité des liens d'amitié déclarée du réseau « R_Friend »
- un seuil d'un appel par mois au minimum, soit 3 appels par vague pour le réseau « R_Call ».

L'intersection des réseaux conduit à un réseau multiplexe de 109 individus permanents. Il comprend 2 volets : les réseaux « R_Friend » et « R_Call » redimensionnés. Le tableau n°14 présente les métriques de base du réseau et les tableaux n°15 et n°16 (p.128, 129) les modélisations du réseau.

Tableau n°14 : évolution des métriques de base pour le réseau multiplexe « R_Friend ∩ R_Call » (seuil sur l'intensité des liens; 109 noeuds)

Vague d'observation **Métriques R_Friend [R_Call]**	**n°1**	**n°2**	**n°3**
Densité	0,074 [0,031]	0,086 [0,033]	0,087 [0,030]
Degré moyen	8,018 [3,376]	9,284 [3,523]	9,394 [3,248]
Nombre de liens	437 [184]	506 [192]	512 [177]
Index de Jaccard	0,693 [0,679]	0,792 [0,670]	

Examen des métriques descriptives

L'intersection produit une faible perte d'effectif d'individus - soit environ 11% - en comparaison avec le réseau de référence « R_Friend ». Le nombre de liens du réseau multiplexe est réduit de 10% en moyenne pour le réseau « R_Friend » résultant et d'environ 3% pour le réseau « R_Call », ce que nous considérons comme négligeable. Les densités et les degrés moyens issus de la combinaison des 2 réseaux uniplexes sont comparables à ceux des réseaux initiaux (tableau n°14). Enfin, les index de Jaccard obtenus sont très proches de ceux observés pour les réseaux uniplexes initiaux (tableau n°14).

Le modèle « naïf » (FC1)

Le modèle « FC1 » comprend uniquement les variables socio-démographiques relatives aux attributs des individus, l'effet de contrôle de densité et les effets structuraux inter-réseaux qui rendent comptent de l'influence mutuelle de l'entrelacement des réseaux. Nous qualifions ce modèle de « naïf » (tableau n°15, p.128).

Le volet « R_Friend » (modèle FC1a)

Les variables socio-démographiques

Pour le volet « R_Friend » du modèle multiplexe naïf (FC1a, tableau n°15), parmi les co-variables seules celles relatives à une homophilie liée à l'origine ethnique blanche (« même ethnie B ») et à l'interaction du sexe et du nombre d'enfants sont statistiquement significatives.

Pour les déclarations d'amitiés, en l'absence de prise en compte d'effets structuraux intra-réseaux, le fait de partager la même origine ethnique blanche a un effet négatif, faiblement significatif : on observe ainsi une évolution faible mais perceptible vers une hétérophilie liée à l'ethnie dans le réseau des relations d'amitiés entrelacé avec celui des interactions à distance.

En comparaison avec les résultats obtenus pour le réseau uniplexe « R_Friend », les effets relatifs à l'ethnie d'Alter et au nombre d'enfants d'Alter n'apportent plus d'explication à l'évolution du réseau des amitiés déclarées considéré dans une configuration multiplexe avec le réseau des contacts téléphoniques.

Les facteurs structuraux multiplexes

L'examen de la significativité des facteurs multiplexes révèle l'importance majeure de l'effet positif direct du réseau des amitiés déclarées sur celui des contacts téléphoniques. L'effet triadique mixte de fermeture entre les 2 réseaux (*closure of mixted two-paths*) est également très significativement positif : il traduit une tendance forte des individus à faire les mêmes choix de relations amicales que les choix amicaux des individus avec lesquels ils ont des échanges distanciels communs. En revanche, il n'y a pas d'évidence statistiquement significative d'une concordance des choix relationnels amicaux des individus compte tenu de leurs choix de relations par téléphone.

Le volet « R_Call » (modèle FC1b)

Les variables socio-démographiques

Pour le volet « R_Call » (FC1b), l'effet « sexe d'Alter » a un effet négatif significatif : la tendance à avoir des interactions téléphoniques avec les individus de sexe féminin[20] est significativement faible. En revanche, les effets relatifs à des homophilies liées au sexe, aux origines ethniques et au nombre d'enfants ne sont pas statistiquement significatifs aux seuils considérés dans notre étude.

Les résultats obtenus pour le réseau uniplexe « R_Call» montraient des effets significatifs relatifs à une homophilie de sexe et au nombre d'enfants d'Alter. Dans la configuration multiplexe avec le réseau des amitiés déclarées, ces facteurs ne sont plus significatifs.

En revanche, le facteur lié au sexe d'Alter qui n'était pas significatif dans la configuration uniplexe du réseau « R_Call » est très significatif dans la version multiplexe. Ainsi, l'homophilie de sexe constatée dans la configuration uniplexe apparaît plutôt se révéler être une popularité relative au sexe lors de l'examen des relations multiplexes.

[20] Pour mémoire la variable sexe est codée comme suit : Homme (0), Femme (1) (cf. chapitre 3).

Les facteurs structuraux multiplexes

L'effet multiplexe direct du réseau des relations distancielles sur les relations amicales, ainsi que l'effet triadique mixte de fermeture sont significatifs avec des coefficients positifs. Toutefois, l'effet direct est nettement plus prégnant que celui relatif à la fermeture multiplexe.

En résumé, le modèle multiplexe naïf « FC1 », c'est-à-dire sans effets structuraux intra-réseaux autre que celui relatif à la densité, met en évidence :

- une tendance de choix hétérophiles dans les nominations de relations amicales pour les individus de même origine ethnique B, mais aucun effet d'homophilie pour les relations à distance.
- un effet négatif significatif de la popularité de la variable « sexe » pour les interactions à distance.
- des effets directs et triadiques mixtes de fermetures multiplexes significatifs.
- l'absence d'effet significatif de concordance des choix entre les 2 réseaux.

<u>Le modèle « complet » (FC2)</u>

Pour le modèle « FC2 », nous incluons les facteurs structuraux intra-réseaux relatifs à : l'activité et popularité des individus, la tendance à l'appariement avec des individus de degré équivalent et la transitivité relationnelle. Nous le qualifions de modèle « complet » (tableau n°16, p.129).

Les facteurs structuraux intra-réseaux

Les 3 facteurs structuraux intra-réseaux introduits ont des effets très significatifs pour les 2 volets du réseau multiplexe. Cependant, des différences apparaissent pour les signes et l'amplitude des effets.

Pour les 2 volets du réseau multiplexe, les effets des facteurs relatifs à la transitivité relationnelle triadique et à l'effet d'activité et de popularité des individus sont très significatifs et positifs. L'effet qui traduit l'appariement préférentiel des individus est également très significatif mais négatif. Pour les deux derniers effets cités, l'hypothèse d'explication du phénomène est similaire à celle donnée précédemment. D'une part, l'évolution vers l'hétérophilie du point de vue de la connectivité à des individus est perceptible dans ce réseau multiplexe. D'autre part, le facteur structural de popularité et d'activité révèle un effet positif sur la dynamique des interactions dans une configuration multiplexe.

Les variables socio-démographiques

Les variables relatives au sexe (volet R_Call), d'une part ainsi qu'au nombre d'enfants d'Alter et à la même appartenance à l'ethnie B (volet R_Friend), d'autre part, sont les seules variables statistiquement significatives.

Pour le volet « R_Friend » (FC2a), les variables « nombre d'enfants d'Alter » et « même ethnie B » ont des seuils de significativité faibles et de signes opposés. En ce qui concerne la variable « même ethnie B », la significativité de l'effet et son amplitude sont très proches de ceux observés pour le modèle « naif » « FC1a » : la tendance aux choix hétérophiles dans les déclarations de relations amicales pour l'ethnie blanche demeure lors de la prise en compte des effet endogènes structuraux.

Pour le volet « R_Call » (FC2b), seule la variable « sexe d'Alter » est significative. Toutefois, en comparaison avec le modèle naïf « FC1b », sa significativité est plus faible et le coefficient, de signe inchangé, est d'ampleur réduite.

Les facteurs structuraux multiplexes

En ce qui concerne les facteurs structuraux multiplexes, les effets directs et triadiques de fermeture demeurent significatifs pour les 2 volets du réseau multiplexe. L'effet direct mutuel des 2 réseaux est inchangé en sens et en amplitude : cet effet est positif et 2 fois plus important dans le sens « R_Call » vers « R_Friend ». Par conséquent, ceci souligne que les appels téléphoniques impliquent de façon préférentielle des relations amicales.

Toutefois, l'introduction des facteurs intra-réseaux modifie le sens et l'amplitude du facteur relatif à l'effet fermeture inter-réseau et ce d'une façon différente pour chacun des volets du réseau multiplexe. D'une part, pour le volet « R_Call » (FC2b) le niveau de significativité de l'effet reste le même mais le sens de l'impact du facteur change, avec un coefficient négatif de valeur absolue plus élevée. D'autre part, pour le volet « R_Friend » (FC2a), le niveau de significativité de l'effet est plus faible. L'ampleur du coefficient est divisée par 2 tout en conservant un signe positif. Ainsi, pour le volet « R_Friend » (FC2a) l'effet de fermeture apparaît en partie « absorbé » par les effets exprimés par les facteurs structuraux intra-réseaux introduits. Il est probable que l'effet relatif aux triades transitives prenne une part du pouvoir explicatif de la transitivité multiplexe.

En revanche, pour le volet « R_Call » (FC2b) l'introduction des facteurs structuraux intra-réseaux inverse le sens donné à la fermeture structurale inter-réseaux : l'effet est maintenant significativement négatif. L'hypothèse explicative est ici plus ardue à concevoir. Nous pouvons toutefois énoncer que la prise en compte du nombre de triades transitives modifie l'explication apportée par la transitivité multiplexe exprimée par la fermeture entre les 2 réseaux. Ceci révèlerait la faible tendance des individus à faire les mêmes choix d'interactions par téléphone que les choix des individus avec lesquels ils ont des liens amicaux communs lorsque la transitivité « intra-réseau » est prise en compte.

Synthèse

Lors de l'introduction des facteurs structuraux intra-réseaux, le modèle multiplexe met en évidence les points saillants suivants :

- la transitivité relationnelle dans les échanges à distance et les nominations d'amis est très significativement positive : la dynamique traduite par la formule « les amis de mes amis sont mes amis » est prégnante dans les déclarations d'amitié et dans les interactions factuelles à distance. La cohérence de la transitivité sur ces 2 dimensions est avérée.
- les individus ont tendance à faire des choix hétérophiles en termes de connectivité dans leurs choix de nomination d'amis et cela se confirme dans les interactions factuelles à distance.
- les effets multiplexes directs et triadiques mixtes de fermeture sont significatifs : l'entrelacement des liens de sociabilité déclarés et effectifs apportent une « plus-value » explicative à celle apportée par les effets structuraux propres à chacun des réseaux. Le récit à 2 voix de la sociabilité éclaire sa dynamique.
- il n'y pas d'évidence significative ni de concordance des choix de nomination d'amis en fonction du choix des interactions à distance, ni du phénomène réciproque.

En ce qui concerne les variables socio-démographiques, dans le cadre de notre problématique nous retiendrons essentiellement que la tendance aux choix hétérophiles dans les déclarations de relations amicales pour l'ethnie blanche demeure lors de la prise en compte des effets structuraux endogènes. En comparaison avec les résultats obtenus avec la modélisation du réseau uniplexe « R_Friend » pour l'origine ethnique blanche, la tendance à l'hétérophilie dans les nominations d'amis bien qu'affaiblie persiste à expliquer l'évolution de ce modèle multiplexe d'ordre 2. Toutefois, ce phénomène n'apparait pas significatif pour expliquer les interactions à distance.

Les effets structuraux multiplexes révèlent l'importance de l'effet direct du réseau des amitiés déclarées sur celui des contacts téléphoniques, cet effet étant par ailleurs plus prégnant en amplitude que dans sa formulation réciproque. Les modèles multiplexes ne montrent pas d'évidence d'une explication de concordance des choix relationnels amicaux des individus entre les 2 réseaux dans l'évolution du réseau multiplexe, cependant il apparaît une tendance très significative des individus à faire les mêmes choix de relations amicales que les choix des individus avec lesquels ils ont des échanges communs à distance.

Bien que les 2 réseaux ne soient pas de même nature, l'un est basé sur des liens déclaratifs et l'autre sur des liens factuels selon une modalité de communication à distance, toutefois leur « logique » relationnelle ne semble ni « étanche » ni « décloisonnée » : la complexité de l'entrelacement des récits de la sociabilité déclarée et factuelle est ici perceptible.

Tableau n°15 : résultats des modélisations longitudinales pour le réseau multiplexe « R_Friend∩R_Call » (109 nœuds)

Réseau Multiplexe « R_Friend∩Call » (seuilé)	**Modèle FC1a**		**Modèle FC2a**	
volet « R_Friend »				
Effets	paramètre	(é.t)	paramètre	(é.t)
R_Friend rate (periode1)	2,188	(0,206)	2,258	(0,206)
R_Friend rate (periode2)	1,343	(0,134)	1,356	(0,134)
Effets structuraux				
R_Friend : degré (densité)	–1,915***	(0,130)	–1,916***	(0,204)
R_Friend : appariement degré[1] équivalent			–0,147***	(0,044)
R_Friend : activité + popularité			0,062***	(0,018)
R_Friend : triades transitives			0,286***	(0,059)
R_Friend : R_Call[(a)]	1,442***	(0,342)	1,314***	(0,327)
R_Friend : R_Call to agreement[(b)]	1,290***	(0,175)	0,556*	(0,222)
R_Friend : from R_Call agreement[(c)]	–0,144	(0,147)	–0,081	(0,124)
Effets des co-variables				
R_Friend : sexe d'Alter	0,181	(0,164)	0,225	(0,166)
R_Friend : même sexe	–0,008	(0,114)	0,014	(0,117)
R_Friend : ethnie_A Alter	–0,121	(0,268)	–0,181	(0,253)
R_Friend : même ethnie_A	0,158	(0,139)	0,119	(0,140)
R_Friend : ethnie_B Alter	0,071	(0,245)	0,098	(0,232)
R_Friend : même ethnie_B	–0,263†	(0,135)	–0,257†	(0,135)
R_Friend : nombre d'enfants d'Alter	0,217	(0,133)	0,223†	(0,124)
R_Friend : similarité du nbre d'enfants	0,062	(0,181)	0,020	(0,179)
R_Friend : inter. sexe*nbre d'enfants	–0,465*	(0,215)	–0,323	(0,227)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence <0,25.
Seuils de significativité statistique : † p<0,1; * p<0,05; ** p<0,01; *** p<0,001; (é.t) : écart-type.
(1) racine carrée des degrés.
(a) R_Friend : R_Call : effet direct du réseau « R_Friend » sur le réseau « R_Call ».
(b) R_Friend : R_Call to agreement : tendance des acteurs à faire les mêmes choix relationnels au sein du réseau « R_Friend » que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau « R_Call ».
(c) R_Friend : from R_Call agreement : concordance des choix relationnels des acteurs au sein du réseau « R_Friend » compte tenu de leurs choix relationnels au sein du réseau « R_Call ».

Tableau n°16 : résultats des modélisations longitudinales pour le réseau multiplexe « R_Friend∩R_Call » (109 nœuds)

Réseau Multiplexe « R_Friend∩Call » (seuil)	**Modèle FC1b**		**Modèle FC2b**	
volet « R_Call »				
Effets	paramètre	(é.t)	paramètre	(é.t)
Effets structuraux				
R_Call rate (periode1)	0,869	(0,117)	0,913	(0,131)
R_Call rate (periode2)	0,875	(0,123)	0,906	(0,127)
R_Call : degré (densité)	–2,936***	(0,251)	–3,364***	(0,424)
R_Call : appariement degré[1] équivalent			–0,230*	(0,091)
R_Call : activité + popularité			0,103***	(0,026)
R_Call : triades transitives			0,980***	(0,178)
R_Call : R_Friend[(a)]	2,014***	(0,323)	2,651***	(0,384)
R_Call : R_Friend to agreement[(b)]	0,191*	(0,077)	–0,351*	(0,150)
R_Call : from R_ Friend agreement[(c)]	–0,009	(0,038)	0,037	(0,039)
Effets des co-variables				
R_Call : sexe d'Alter	–0,806**	(0,299)	–0,602*	(0,283)
R_Call : même sexe	0,049	(0,179)	0,063	(0,195)
R_Call : ethnie_A Alter	0,129	(0,411)	0,013	(0,417)
R_Call : même ethnie_A	0,205	(0,219)	0,173	(0,233)
R_Call : ethnie_B Alter	–0,366	(0,417)	–0,733	(0,454)
R_Call : même ethnie_B	–0,253	(0,236)	–0,278	(0,255)
R_Call : nombre d'enfants d'Alter	0,016	(0,242)	–0,007	(0,265)
R_Call : similarité du nbre d'enfants	–0,022	(0,293)	–0,077	(0,320)
R_Call : inter. sexe*nbre d'enfants	0,162	(0,404)	–0,116	(0,480)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence <0,25.
Seuils de significativité statistique : † p<0,1; * p<0,05; ** p<0,01; *** p<0,001; (é.t) : écart-type.
(1) racine carrée des degrés.
(a) R_Call : R_Friend : effet direct du réseau « R_Call » sur le réseau « R_Friend ».
(b) R_Call : R_Friend : to agreement : tendance des acteurs à faire les mêmes choix relationnels au sein du réseau « R_Call » que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau « R_Friend ».
(c) R_Call : from R_Friend agreement : concordance des choix relationnels des acteurs au sein du réseau « R_Call » compte tenu de leurs choix relationnels au sein du réseau « R_Friend ».

4.2.5 Modèles du réseau « R_Friend ∩ R_Prox »

La deuxième modélisation multiplexe est réalisée à partir de la combinaison « R_Friend ∩ R_Prox ». Les seuils appliqués sur l'intensité des liens sont les suivants :

- une valeur minimale de 3 pour l'intensité d'amitié déclarée pour le réseau « R_Friend ».
- une heure minimum de co-présence par semaine pour le réseau « R_ Prox ».

L'intersection des deux réseaux conduit à un réseau multiplexe de 111 individus. Il comprend 2 volets constitués par les réseaux « R_Friend » et « R_Prox » redimensionnés. Le tableau n°17 présente les métriques de base de ce réseau multiplexe.

Tableau n°17 : évolution des métriques du réseau multiplexe « R_Friend ∩ R_Prox » (seuillage de l'intensité des liens ; 111 nœuds)

Vague d'observation **Métriques R_Friend [R_Prox]**	**n°1**	**n°2**	**n°3**
Densité	0,072 [0,051]	0,083 [0,062]	0.085 [0,061]
Degré moyen	7,874 [5,586]	9,171 [6,775]	9.387 [6,685]
Nombre de liens	437 [310]	509 [376]	521 [371]
Index de Jaccard	0,686 [0,711]	0,801 [0,762]	

La perte d'effectif qui résulte de la combinaison des 2 réseaux est faible (environ 9%) en comparaison avec le réseau de référence « R_Friend ». Le nombre de liens du réseau multiplexe est réduit de 9% en moyenne pour le réseau « R_Friend » résultant, et d'environ 7% pour le réseau « R_Prox ». Nous considérons cette attrition négligeable. Les densités et les degrés moyens résultants de la combinaison des 2 réseaux uniplexes sont comparables à ceux des réseaux initiaux. Enfin, les index de Jaccard obtenus sont proches de ceux observés pour les réseaux uniplexes initiaux.

Les tableaux n°18 et n°19 suivants présentent les résultats des modélisations longitudinales du réseau multiplexe « R_Friend ∩ R_Prox ».

Tableau n°18 : résultats des modélisations longitudinales pour le réseau multiplexe « R_Friend∩R_Prox » (seuillage ; 111 nœuds)

Réseau Multiplexe « R_Friend∩R_Prox » (seuil)	**Modèle FP1a**		**Modèle FP2a**	
Effets	paramètre	(é.t)	paramètre	(é.t)
Volet « R_Friend »				
R_Friend rate (period1)	1,955	(0,168)	2,212	(0,221)
R_Friend rate (period2)	1,152	(0,118)	1,234	(0,130)
Effets structuraux				
R_Friend : degré (densité)	–2,100***	(0,146)	–2,080***	(0,207)
R_Friend : appariement degré[1] équivalent			–0,159***	(0,043)
R_Friend : activité + popularité			0,058***	(0,017)
R_Friend : triades transitives			0,421***	(0,041)
R_Friend : R_Prox[(a)]	1,199***	(0,314)	1,466***	(0,345)
R_Friend : R_Prox to agreement[(b)]	0,802***	(0,081)	0,300***	(0,088)
R_Friend : from R_Prox agreement[(c)]	-0,110	(0,071)	–0,091	(0,070)
Effets des co-variables				
R_Friend : sexe d'Alter	–0,454*	(0,179)	–0,354*	(0,165)
R_Friend : même sexe	0,183	(0,121)	0,186	(0,119)
R_Friend : ethnie_A Alter	0,558†	(0,295)	0,208	(0,287)
R_Friend : même ethnie_A	0,200	(0,147)	0,123	(0,145)
R_Friend : ethnie_B Alter	1,041***	(0,304)	0,776**	(0,257)
R_Friend : même ethnie_B	–0,096	(0,151)	–0,061	(0,143)
R_Friend : nombre d'enfants d'Alter	0,154	(0,130)	0,137	(0,130)
R_Friend : similarité du nbre d'enfants	0,068	(0,189)	0,091	(0,179)
R_Friend : inter. sexe*nbre d'enfants	–0,435*	(0,222)	–0,343	(0,229)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence <0,25.
Seuils de significativité statistique : † p<0,1; * p<0,05; ** p<0,01; *** p<0,001; (é.t) : écart-type.
(1) racine carrée des degrés.
(a) R_Friend : R_Prox : effet direct du réseau « R_Friend » sur le réseau « R_Prox ».
(b) R_Friend : R_Prox to agreement : tendance des acteurs à faire les mêmes choix relationnels au sein du réseau « R_Friend ».
que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau « R_Prox ».
(c) R_Friend : from R_Prox agreement : concordance des choix relationnels des acteurs au sein du réseau « R_Friend » compte tenu de leurs choix relationnels au sein du réseau « R_Prox ».

Tableau n°19 : résultats des modélisations longitudinales pour le réseau multiplexe « R_Friend∩R_Prox » (seuillage ; 111 nœuds)

Réseau Multiplexe « R_Friend∩R_Prox » (seuil)	**Modèle FP1b**		**Modèle FP2b**	
Effets	paramètre	(é.t)	paramètre	(é.t)
Volet « R_Prox »				
R_Prox rate (period1)	1,251	(0,128)	1,814	(0,231)
R_Prox rate (period2)	1,055	(0,115)	1,370	(0,163)
Effets structuraux				
R_Prox : degré (densité)	–1,815***	(0,154)	–0,900**	(0,282)
R_Prox : appariement degré[1] équivalent			–0,295*	(0,126)
R_Prox : activité + popularité			0,046	(0,055)
R_Prox : triades transitives			0,921***	(0,060)
R_Prox : R_Friend[(a)]	1,006***	(0,276)	1,098***	(0,290)
R_Prox : R_Friend to agreement[(b)]	0,604***	(0,083)	0,048	(0,105)
R_Prox : from R_Friend agreement[(c)]	–0,116**	(0,037)	–0,080*	(0,039)
Effets des co-variables				
R_Prox : sexe d'Alter	–0,104	(0,193)	0,210	(0,202)
R_Prox : même sexe	–0,102	(0,139)	–0,123	(0,136)
R_Prox : ethnie_A Alter	0,302	(0,292)	0,510	(0,315)
R_Prox : même ethnie_A	0,066	(0,158)	–0,071	(0,164)
R_Prox : ethnie_B Alter	0,070	(0,293)	0,160	(0,303)
R_Prox : même ethnie_B	–0,014	(0,160)	0,093	(0,158)
R_Prox : nombre d'enfants d'Alter	0,051	(0,153)	0,155	(0,166)
R_Prox : similarité du nbre d'enfants	–0,309	(0,221)	–0,336	(0,220)
R_Prox : inter. sexe*nbre d'enfants	0,322	(0,238)	0,084	(0,272)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence <0,25.
Seuils de significativité statistique : † p<0,1; * p<0,05; ** p<0,01; *** p<0,001; (é.t) : écart-type.
(1) racine carrée des degrés.
(a) R_Prox : R_Friend : effet direct du réseau « R_Prox » sur le réseau « R_Friend ».
(b) R_Prox : R_Friend : to agreement : tendance des acteurs à faire les mêmes choix relationnels au sein du réseau « R_Prox » que les choix des acteurs avec lesquels ils ont une relation commune au sein du réseau « R_Friend ».
(c) R_Prox : from R_Friend agreement : concordance des choix relationnels des acteurs au sein du réseau « R_Prox » compte tenu de leurs choix relationnels au sein du réseau « R_Friend ».

Le modèle « naïf » (FP1)

Le modèle « FP1 » (tableau n°18, p.131) comprend uniquement les variables socio-démographiques relatives aux attributs des individus, l'effet lié à la densité des réseaux et les effets structuraux inter-réseaux qui rendent comptent de l'influence mutuelle des réseaux. Nous le qualifions de modèle « naïf ».

Les variables socio-démographiques

Pour le volet « R_Friend », les facteurs relatifs aux variables « sexe d'Alter », aux origines ethniques d'Alter (A et B) et l'interaction du sexe avec le nombre d'enfants sont significatifs à des seuils de significativité très contrastés. Plus précisément, l'origine ethnique B d'Alter est le facteur dont l'effet est le plus significatif parmi les variables socio-démographiques, avec un coefficient positif. Le facteur « origine ethnique A d'Alter » est peu significatif, sans être toutefois négligeable, avec un coefficient positif. Par conséquent, on observe ici un effet largement dominant de popularité dans les choix préférentiels de nominations de liens d'amitiés en significativité pour les individus d'origine ethnique blanche.

Pour le volet « R_Prox », aucune variable socio-démographique n'est statistiquement significative aux seuils considérés dans cette étude. On n'observe aucun effet de tendance homophile ou hétérophile dans les co-présences. Nous faisons l'hypothèse que ce constat peut être expliqué par la structure, dans une dimension spatiale et organisationnelle. En effet, d'une part, les individus sont logés dans une même résidence ce qui crée des opportunités de rencontres aléatoires qui ne relèvent pas d'un choix homophile ou hétérophile dans la rencontre physique telle que détectée. D'autre part, des activités communes sont régulièrement organisées par la résidence familiale du MIT. La participation des individus à ces activités de groupe créée une détection de co-présence qui n'est pas forcément la résultante d'un choix d'appariement dyadique.

En l'absence d'informations sur la disposition spatiale des logements et le contenu et la forme des évènements de groupe, nous ne pouvons pas affirmer une tendance de recherche de contacts homophiles ou hétérophiles. La prise en compte des co-présences permet de prendre en compte un effet relatif à la structure d'opportunités dans l'explication de l'évolution du réseau. Cet effet traduit l'impact de l'organisation spatiale des logements, des activités de groupe et des activités dyadiques choisies mais nous ne pouvons pas séparer leurs effets propres.

Les facteurs structuraux multiplexes

Pour les 2 volets du réseau multiplexe, l'effet direct « R_Friend » vers « R_Prox » et l'effet réciproque sont statistiquement très significatifs et positifs. De même, la tendance des individus à faire les mêmes choix relationnels au sein du réseau « R_Friend » que les choix des individus avec qui ils sont co-présents est également très significative, ainsi que l'effet réciproque.

En ce qui concerne le volet « R_Prox », le facteur structural dyadique mixte relatif à la concordance des choix relationnels est significatif. Le signe du coefficient facteur est négatif : par conséquent la concordance des nominations d'amis et des détections des co-présences explique négativement l'évolution du réseau. Ceci tend à confirmer l'impact des co-présences non choisies au niveau dyadique (les activités de groupes et la répartition spatiale des logements) dans l'évolution du réseau.

Le modèle « complet » (FP2)

Le modèle « FP2 » (tableau n°19, p.132) comprend, en plus des variables exogènes et des facteurs endogènes multiplexes inclus dans le modèle « FP1 », les facteurs structuraux intra-réseaux relatifs à : l'activité et popularité des individus, la tendance à l'appariement avec des individus de degré équivalent et la transitivité relationnelle.

Les facteurs structuraux intra-réseaux

Parmi les 3 effets structuraux intra-réseaux introduits, celui qui traduit la transitivité relationnelle est statistiquement très significatif pour les 2 volets du réseau multiplexe, ce qui confirme les mêmes tendances transitives observées et commentées précédemment. Le facteur relatif à l'activité et à la popularité des individus a un effet statistiquement très significatif et positif pour le volet « R_Friend » (FP2a). Toutefois, on n'observe pas d'effet de la popularité structurale pour les co-présences (modèle FP2b). Enfin, l'effet qui traduit la tendance des individus à se lier préférentiellement avec des pairs de degré équivalent est significativement négatif pour les 2 volets du réseau multiplexe. On observe ainsi une tendance significative d'appariement dissimilaire du point de vue de la connectivité dans les amitiés déclarées et les co-présences détectées.

L'introduction des facteurs intra-réseaux modifie sensiblement, d'une part, les effets des facteurs qui traduisent les phénomènes d'une homophilie liée aux attributs des individus et, d'autre part, les effets des facteurs structuraux triadiques mixtes inter-réseaux.

L'évolution des effets des variables socio-démographiques

Pour le volet « R_Friend » (modèle FP2a) seules 2 variables socio-démographiques conservent des effets significatifs : le sexe d'Alter et l'origine ethnique B d'Alter. L'introduction des éléments structuraux intra-réseaux réduit la significativité de leurs effets respectifs et l'amplitude des coefficients de leurs facteurs, mais les choix de nominations demeurent significativement et positivement expliqués par la popularité des individus d'origine ethnique blanche et négativement par la modalité « femme » du sexe des individus.

En revanche, les facteurs relatifs à « l'origine ethnique A d'Alter » et à l'interaction du sexe des individus avec le nombre d'enfants ne sont plus statistiquement significatifs dans ce modèle complet : les effets structuraux introduits semblent médiatiser ces effets.
Pour le volet « R_Prox » (FP2b), l'introduction des facteurs structuraux intra-réseaux ne change pas le constat établi précédemment : nous n'observons aucun effet ni de tendance homophile ou hétérophile, ni de popularité d'un attribut dans les échanges de proximité entre les individus.

L'évolution des effets structuraux inter-réseaux

Pour les 2 volets du réseau multiplexe, les effets directs d'un réseau sur l'autre sont statistiquement très significatifs.

D'une part, pour le volet « R_Friend » (FP2a), l'introduction des facteurs structuraux intra-réseaux ne change pas la haute significativité de l'effet positif qui traduit la tendance des individus à faire les mêmes choix relationnels au sein du réseau « R_Friend » que les choix des individus avec lesquels ils sont en co-présence. Cependant, comme nous en avons déjà fait l'hypothèse, là encore l'introduction du facteur relatif au nombre de triades transitives réduit l'ampleur de l'effet de fermeture structurale dyadique mixte. Enfin, comme cela a été observé précédemment, le facteur relatif à la concordance des choix relationnels n'est pas statistiquement significatif pour ce volet du réseau multiplexe.

D'autre part, pour le volet « R_Prox » (FP2b) l'introduction des facteurs structuraux intra-réseaux conduit à la perte de significativité de la tendance des individus à être co-présents avec les individus avec qui ils ont une relation commune au sein du réseau « R_Friend ».

Par ailleurs, l'effet relatif à la concordance perd un peu de sa significativité : son facteur conserve son signe négatif, mais son coefficient est réduit de plus d'un tiers en valeur absolue en comparaison avec le même effet dans le modèle naïf. Ainsi, pour le volet « R_Prox » (FP2b), l'explication portée par ce phénomène dans la dynamique relationnelle de proximité entre les individus apparaît en partie « absorbée » par les facteurs structuraux intra-réseaux introduits.

Synthèse

Les trois effets multiplexes sont significatifs ce qui souligne la force de l'entrelacement des liens déclaratifs d'amitié et de la co-présence des individus. Toutefois, leur interprétation est complexe.

Le réseau « R_Prox », bâti sur la détection des co-présences des individus, médiatise un potentiel effet de structure d'opportunité. En effet, cette détection informe que des individus, deux ou plus, sont à un moment donné et de façon répétée dans un espace commun et à faible distance (sociale[21]). Cette situation spatiale agrège la possibilité matérielle d'échanges dyadiques en face-à-face, mais également une présence simultanée qui peut ne pas donner lieu à un échange verbal ou non verbal. En effet, d'une part, les individus sont logés dans une même résidence ce qui crée des opportunités de rencontres aléatoires qui ne relèvent pas d'un choix homophile ou hétérophile dans la rencontre physique telle que détectée. D'autre part, des activités communes sont régulièrement organisées par la résidence familiale du MIT. La participation des individus à ces activités de groupe créée une détection de co-présence qui n'est pas forcément la résultante d'un choix de relation dyadique. En l'absence d'informations sur la disposition spatiale des logements et le contenu et la forme des évènements de groupe, nous ne pouvons pas affirmer une tendance de recherche de contacts homophiles ou hétérophiles. Nous faisons l'hypothèse que cette présence simultanée, compte-tenu du seuil horaire de co-présence appliqué, peut être en grande partie expliquée par l'organisation des nombreux évènements par l'institution qui gère la résidence communautaire, ces évènements constituant des foyers d'activité. Dans ce cas, nous pouvons considérer que le réseau « R_Prox », à travers ses effets dans la configuration multiplexe, médiatise une partie des effets relatifs à la structure d'opportunités. Ces effets comportent l'impact de l'organisation spatiale des logements, des activités de groupe et des activités dyadiques choisies. Toutefois, dans cette expérimentation nous ne pouvons pas discerner leurs parts propres respectives.

[21]Toute ressemblance avec la distance sociale préconisée en situation pandémique est (relativement) fortuite.

4.2.6 Le réseau « R_Friend ∩ R_Call ∩ R_Prox »

La dernière modélisation multiplexe est conduite à partir de la combinaison d'ordre 3 des réseaux « R_Friend ∩ R_Call ∩ R_Prox ». Les seuils appliqués sur l'intensité des liens sont les suivants :

- une valeur minimale de 3 pour l'intensité d'amitié déclarée pour le réseau « R_Friend ».
- 1 appel par mois au minimum pour le réseau « R_Call ».
- 1 heure minimum de co-présence par semaine pour le réseau « R_ Prox ».

L'intersection des deux réseaux conduit à un réseau multiplexe de 104 individus. Le tableau n°20 présente les métriques de base du réseau multiplexe.

Tableau n°20 : évolution des métriques de base pour le réseau multiplexe d'ordre 3 « R_Friend ∩ R_Call ∩ R_Prox » (seuillage; 104 noeuds)

Vague d'observation Métriques R_Friend [R_Call] { R_Prox }	**n°1**	**n°2**	**n°3**
Densité	0,082 [0,035] {0,061}	0,094 [0,035] {0,070}	0.097 [0,032] {0,070}
Degré moyen	8,442 [3,558] {6,327}	9,654 [3,635] {7,250}	9.942 [3,346] {7,231}
Nombre de liens	439 [185] {329}	502 [189] {377}	517 [174] {376}
Index de Jaccard	0,702 [0,670] {0,726}		0,800 [0,681] {0,759}

Examen des métriques du réseau multiplexe d'ordre 3

La combinaison des 3 réseaux uniplexes produit un réseau multiplexe qui comprend un nombre d'individus réduit de 15% en comparaison avec le réseau de référence « R_Friend ». Cette attrition est la plus importante des réseaux multiplexes produits. Toutefois, elle demeure acceptable[22].

Cependant, l'examen des autres métriques (tableau n°20) révèle que pour chaque volet du réseau multiplexe les densités et les degrés moyens sont plus importants qu'au sein des réseaux uniplexes respectifs. Ainsi, bien que le nombre global de liens soit en baisse, le réseau multiplexe d'ordre 3 est d'une connectivité plus importante que les réseaux uniplexes : sa cohésion globale est plus importante et les individus qui le composent sont en moyenne plus actifs. Les index de Jaccard sont d'un niveau équivalent à ceux observés précédemment.

[22] Le seuil limite d'attrition que nous fixons est de 20%. Ce seuil correspond à la valeur préconisée par Ripley *et al.*, (2019, p.33).

Compte-tenu de la longueur des tableaux qui présentent les résultats des modélisations du réseau multiplexe d'ordre 3, ceux-ci sont reportés en intégralité en Annexe (tableau A.12, p.173). Nous exposons ci-dessous les éléments saillants que révèlent la modélisation multiplexe d'ordre 3 en comparaison avec les modélisations commentées précédemment.

Le modèle « naïf » (FCP1)

Les variables socio-démographiques

Quatre variables socio-démographiques ont des effets statistiquement significatifs et ce uniquement pour les volets « R_Friend » et « R_Call », avec des seuils de significativité statistique et des signes différents :

- les effets de popularité relatifs aux 2 origines ethniques d'Alter (A et B), ainsi que le nombre d'enfants d'Alter (volet « R_Friend », modèle FCP1a,) sont positifs et hautement significatifs.
- l'effet d'homophilie de sexe (volet « R_Call », modèle FCP1b,) est significativement négatif.

Ces deux effets étaient identifiés par les modélisations uniplexes, mais ils n'apparaissaient pas significatifs dans le modèle multiplexe d'ordre 2 « R_Friend∩R_Call ». La modélisation multiplexe d'ordre 3 confirme les effets de ces facteurs dans une configuration qui introduit le réseau des co-présences.

Pour les facteurs inter-structuraux, nous observons :

- la tendance des individus à faire les mêmes choix de nominations d'amis que les acteurs avec lesquels, d'une part, ils ont des interactions à distance et, d'autre part, ils sont en co-présence.
- la tendance des individus à faire les mêmes choix de contacts téléphoniques que les acteurs avec lesquels ils sont en co-présence, ainsi que l'effet réciproque. La force de l'effet est 5 fois plus élevée dans le sens « R_Prox » vers « R_Call » que dans le sens réciproque.
- les effets négatifs de la concordance des co-présences des acteurs compte-tenu de leurs choix relationnels au sein du réseau « R_Friend ».
- des effets directs mutuels très significatifs et positifs entre les réseaux « R_Friend » et « R_Call », d'une part, et « R_Friend » et « R_Prox », d'autre part.

Ces trois effets étaient déjà identifiés en configuration multiplexe d'ordre 2. La modélisation multiplexe d'ordre 3 les confirme.

Le modèle complet (FCP2)

Pour tous les facteurs structuraux intra-réseaux introduits et les volets du réseau multiplexe d'ordre 3, nous soulignons les éléments saillants suivants :

- l'effet du nombre de triades transitives est toujours très significatif et positif.
- l'effet relatif à l'appariement préférentiel des individus avec des pairs de degré équivalent est significativement négatif pour le réseau des relations amicales et des contacts téléphoniques. Cet effet a été constaté pour la modélisation d'ordre 2 entre ces 2 réseaux. Cependant, au sein du réseau des co-présences on n'observe plus aucun effet significatif d'appariement préférentiel en fonction des degrés dans cette configuration multiplexe d'ordre 3. Or, cet effet était significatif dans la modélisation d'ordre 2 « R_Friend∩R_Prox ».
- l'effet relatif à l'activité et à la popularité des individus est positif. Il est également très significatif pour le réseau des échanges téléphoniques et, d'autre part, peu significatif pour le réseau des liens amicaux. Pour le réseau des co-présences cet effet n'est pas statistiquement significatif aux seuils statistiques considérés, comme cela a été constaté pour la modélisation d'ordre 2.

L'introduction des facteurs structuraux intra-réseaux fait évoluer différemment les effets des variables socio-démographiques qui étaient significatives dans le modèle naïf :

- pour le volet « R_Friend » (FCP2a) : on n'observe plus d'effet significatif d'une homophilie relative à l'origine ethnique. Toutefois, on note toujours l'importante significativité des effets positifs de la popularité des origines ethniques A et B, et du nombre d'enfants d'Alter.
- pour le volet « R_Call » (FCP2b) : l'effet négatif de l'homophilie de sexe augmente sensiblement en significativité.
- pour le volet « R_Prox » (FCP2c) : l'effet dû à l'interaction du sexe d'Alter et du nombre d'enfants d'Alter croît en significativité positive.

Enfin, pour les facteurs structuraux spécifiquement multiplexes, les changements les plus notables concernent l'entrelacement du réseau des échanges à distance avec celui des relations d'amitié et celui des co-présences :

- avec l'introduction des facteurs structuraux intra-réseaux, la tendance des individus à faire les mêmes choix relationnels au sein de réseau d'échange par téléphone que les acteurs avec lesquels ils ont une relation commune au sein du réseau d'amitié devient très significative, avec un coefficient de signe négatif. L'effet réciproque demeure très significativement positif, comme dans le modèle naïf.

5. Discussion conclusive

Dans ce chapitre conclusif, nous confrontons les cinq hypothèses de notre problématique aux résultats empiriques obtenus par l'analyse des six réseaux sociaux uniplexes et multiplexes formés à partir de la base relationnelle exploitée. Ensuite, nous discutons des limites méthodologiques de cette étude et de l'approche multiplexe pour appréhender l'entrelacement du récit multidimensionnel de la sociabilité. Nous soulignons la place de notre étude au sein de travaux proches récents et plus anciens. Enfin, nous ouvrons des perspectives sur de possibles suites de cette recherche.

Dans cette étude nous avons présenté l'analyse de plusieurs jeux de données relationnelles sur les interactions entre une centaine de jeunes professionnels en reprise d'études. Ces données ont été collectées par trois sources différentes : des questionnaires sur les relations amicales, des relevés d'échanges par téléphone ainsi qu'une détection des co-présences des individus dans un espace proche susceptible de produire des interactions en face-à-face. Cette collecte de données relationnelles sur un temps long, issues d'une importante expérimentation conduit au sein du MIT (Aharony *et al.*, 2011), nous a permis de conduire des analyses longitudinales de plusieurs réseaux sociaux complets, uniplexes et multiplexes. Nous avons pu dégager des complémentarités, des recouvrements, des assonances mais également des dissonances dans le récit de la sociabilité en examinant séparément et ensemble les relations restituées et effectives, distancielles et présentielles, synchrones et asynchrones.

Notre étude s'inscrit dans un contexte spécifique. En effet, l'université où évoluent les individus suivis est très prestigieuse, elle jouit d'une renommée mondiale et attire des étudiants d'origines ethniques variées. Ainsi, la communauté que nous avons suivie comprend huit origines ethniques différentes. De surcroît, cette « Université-Monde » a une longue tradition d'accueil d'étudiants en famille au sein d'une résidence nommée « *MIT Westgate Family Housing Community* », dont un des atouts promus est : « *Westgate is a beautiful community of friendly and warm people.* ***More importantly, it is a place of long-lasting friendships***. »[23]. Par conséquent, nous avons choisi, d'une part, d'observer l'évolution des relations amicales d'individus représentatifs de cette communauté et, d'autre part, d'apprécier les impacts relatifs et concomitants de l'homophilie exogène au réseau social due à un attribut hérité (l'origine ethnique des étudiants) et de l'homophilie structurale endogène.

---

[23] « *Westgate Guide For Spouses, p.1* » (https://studentlife.mit.edu/housing/graduate-family-housing/graduate-residences/westgate-apartments)

Pour cela, nous avons tout d'abord réalisé des analyses structurales descriptives puis inférentielles des réseaux sociaux. Nous confrontons ci-après nos résultats à nos hypothèses initiales.

L'apport des analyses structurales descriptives

Les analyses structurales descriptives de la dynamique d'évolution des 3 réseaux uniplexes révèlent des points saillants synthétisés dans le tableau n°21.

Tout d'abord, **le réseau des relations d'amitiés déclarées** ne comprend plus aucun individu isolé à la fin du suivi : tous les étudiants sont en lien (non dirigé) de connaissance avec au moins un pair. La cohésion du réseau des amitiés croît et se centralise au fil du temps, que ce soit pour les relations d'amitiés affirmées ou pour l'ensemble du réseau des connaissances. Ainsi, l'analyse de la sociabilité de la communauté restitue des relations dynamiques dans le temps, avec une connectivité qui progresse. L'évolution comparée des indicateurs de transitivité relationnelle fait apparaître que les triades locales d'amis sont prégnantes et stables au cours du temps. L'évolution de la transitivité globale tend à montrer un certain « déverrouillage » de la fermeture structurale relationnelle du réseau considéré dans son ensemble. Les analyses descriptives permettent de discerner qu'un petit nombre d'étudiants d'origine ethnique asiatique, blanche et hispanique se positionnent de façon persistante comme des intermédiaires clefs de connecteurs de sous-groupes distants d'un point de vue structural. La proportion de ces « *gatekeepers* » selon leur origine ethnique est largement supérieure à la proportion des effectifs respectifs de ces trois ethnies au sein de la communauté prise dans son ensemble, notamment pour les étudiants d'origine hispanique. Toutefois, la force de leur position structurale d'intermédiaire décroît au fil du temps : des trous structuraux (Granovetter, 1973) se comblent, le capital social encastré circule plus fluidement.

Tableau n°21 : comparaisons de l'évolution de descripteurs clefs des 3 réseaux uniplexes.

	R_Friend	**R_Call**	**R_Prox**
Cohésion	↗	→	↗
Centralisation	↗	↘	↘
Transitivité globale	↘	↗	↘
Transitivité locale	→	↘	↘
***Gatekeepers*[1] par origine ethnique**	**asiatique, blanche** hispanique	**hispanique**, asiatique, blanche	**asiatique, blanche**
Evolution du rôle des *gatekeepers*	↘	↘	↗

(1) : **en gras** : proportion déséquilibrée au regard des effectifs de l'origine ethnique au sein de la communauté.

Le réseau des échanges à distance montre une sociabilité effective qui contraste avec la sociabilité déclarée. Ce réseau a une cohésion relativement stable au fil du temps, il se décentralise peu à peu et son organisation évolue d'une fermeture relationnelle locale à une plus grande ouverture. Deux groupes homogènes se distinguent autour des deux origines ethniques dominantes en effectif (blanche et asiatique), avec une connectivité au sein des groupes plus élevée que la connectivité entre les groupes. La connexion entre ces deux groupes est assurée par des étudiants d'une origine ethnique minoritaire (hispanique).

Le réseau des co-présences se densifie au cours du temps. Cette évolution est très probablement imputable, d'une part, à une augmentation des interactions dyadiques en face-à-face mais également, d'autre part, à des regroupements d'individus qui partagent un même espace au cours des nombreuses propositions d'activités partagées proposées par la résidence communautaire « *Wesgate* »[24]. Le réseau des co-présences se décentralise avec le temps. Les co-présences inter-groupes se multiplient entre des individus fortement connectés au sein de leurs groupes respectifs. Le rôle des étudiants d'origines ethniques asiatiques et blanches se confirme et devient de plus en plus important. Ceci contraste avec l'évolution observée de leur rôle au sein du réseau des amitiés déclarées et des contacts téléphoniques.

Hypothèses et résultats empiriques

Nous avons formulé quatre hypothèses concernant le pouvoir explicatif comparé de l'homophilie attribuée aux facteurs socio-démographiques et structuraux.

1. hypothèse H_homo : les effets liés à la similarité des attributs socio-démographiques diminuent nettement lors de la prise en compte des phénomènes structuraux endogènes.

Les facteurs choisis pour traduire les phénomènes structuraux endogènes sont les suivants :

2. hypothèse H_s1 : l'activité et la popularité structurale des individus.
3. hypothèse H_s2 : l'appariement avec des pairs structuralement dissimilaires.
4. hypothèse H_s3 : la transitivité relationnelle entre les individus.

Le tableau n°22 synthétise l'information pertinente au regard de nos hypothèses. L'examen de l'évolution des différents effets relatifs à l'homophilie nous permet de valider partiellement notre hypothèse H_homo. En effet, lors de la prise en compte des facteurs structuraux endogènes au sein des réseaux, les effets de similarité des origines ethniques des individus perdent en significativité. En revanche, l'homophilie liée au sexe révélée par le réseau des contacts téléphoniques demeure, voire augmente.

[24] Le lecteur pourra constater la grande variété des activités proposées aux étudiants du MIT de la résidence « *Wesgate familly housing* » sur le site dédié : http://westgate.mit.edu/.

Tableau n°22 : synthèse des effets endogènes et exogènes aux réseaux uniplexes et multiplexes.

Effets des facteurs exogènes et endogènes		Homophilie Attributs socio-démographiques	Homophilie structurale	Popularité Attributs socio-démographiques d'Alter	Popularité/activité structurale	Transitivité (nombre de triades transitives)
Réseaux uniplexes	Sans facteurs endogènes	R_F (- ethnie B) R_C (- sexe F)		R_F (+ ethnie A,B) R_P (+ ethnie A,B) R_F (+ nbre enfants) R_C (+ nbre enfants)		
	Avec facteurs endogènes	R_F (- ethnie B) R_C (- sexe F)	R_F, R_C, R_P (- appariement de degré équivalent)	R_F (+ ethnie A,B) R_P (+ ethnie A) R_F (+ nbre enfants) R_C (+ nbre enfants)	R_F (+) R_C (+) R_P (-)	R_F (+) R_C (+) R_P (+)
Réseaux multiplexes d'ordre 2	Sans facteurs endogènes	R_F∩C (- ethnie B)		R_F∩P (+ ethnie A,B) R_F∩P (- sexe)		
	Avec facteurs endogènes	R_F∩C (- ethnie B) R_F∩C (- sexe F)	R_F∩C , R_F∩P (- appariement de degré équivalent)	R_F∩P (+ ethnie A,B) R_F∩P (- sexe) R_F∩C (+ nbr enfants)	R_F∩C (+) R_F∩P (+)	R_F∩C (+) R_F∩P (+)
Réseaux multiplexes d'ordre 3	Sans facteurs endogènes	R_F∩C∩P (- sexe F)		R_F∩C∩P (+ ethnie A,B) R_F∩C∩P (+ nbr enfants)		
	Avec facteurs endogènes	R_F∩C∩P (- sexe F)	R_F∩C∩P (- appariement de degré équivalent)	R_F∩C∩P (+ nbr enfants)	R_F∩C∩P (+)	R_F∩C∩P (+)

Les trois facteurs structuraux endogènes choisis contribuent nettement à expliquer l'évolution des réseaux de nominations d'amis, d'échanges par téléphone et des co-présences. Ceci valide nos trois hypothèses H_s1, H_s2 et H_s3. Toutefois, ces facteurs expliquent l'évolution des réseaux différemment.

D'une part, l'hétérogénéité structurale identifiée au sein des modèles uniplexes et multiplexes exprime une tendance à une évolution préférentielle des liens entre des individus de connectivités dissimilaires : avec le temps, les individus très connectés ont de plus en plus tendance à nommer, à contacter et à être en présence de pairs moins connectés qu'eux même et inversement. Le phénomène se retrouve de façon cohérente dans toutes les configurations simulées. De même, le facteur qui traduit la transitivité relationnelle est observé dans toutes les configurations : que ce soit pour les nominations d'amis, les interactions à distance par téléphone ou les co-présences, la transitivité « les amis de mes amis sont mes amis » est très prégnante. La formation des liens est fortement structurée par la fermeture triadique. Celle-ci, nous le supposons, est déterminée par une combinaison de choix et de préférences individuelles, de proximité structurale mais également d'une fermeture locale (*focal closure*) qui relève de contraintes sur les opportunités de choix parmi les pairs et les contraintes induites par les foyers d'activités fréquentés.

D'autre part, en ce qui concerne le facteur combiné de popularité/activité structurale des individus, celui-ci joue un rôle explicatif positif sauf si on isole le réseau des co-présences au sein duquel cette caractéristique structurale combinée à un effet négatif. L'interprétation est ici plus complexe à dénouer et soulève des questions d'ordre méthodologique que nous développons par la suite.

En effet, la détection des co-présences agrège des relations dyadiques ainsi que des proximités spatiales entre des individus qui ne relèvent pas *forcément* d'interactions entre eux, mais qui participent à un même évènement ou se trouvent opportunément à proximité les uns des autres. Par conséquent, ce réseau agrège probablement des interactions dyadiques et une information sur la participation d'individus à des foyers d'activités communs. Cette intrication ne nous permet pas de discerner les effets propres aux interactions en face-à-face de ceux qui relèvent des opportunités offertes aux individus de participer à des foyers d'activités communs en même temps, dont nous avons souligné l'impact potentiel sur les relations sociales (l'homophilie induite par la structure d'opportunités).

En conclusion, les simulations longitudinales confirment que la seule prise en compte des marqueurs de l'homophilie exogène au réseau (expliquée par les attributs socio-démographiques des individus) nous aurait conduit à sur-valoriser l'importance de ce phénomène, à l'exception toutefois de l'homophilie liée au sexe. Ces observations contribuent à étayer l'hypothèse d'une recherche de relations d'une plus grande variété d'un point de vue structural au dépend d'une homophilie attendue due aux attributs hérités (McPherson et Smith-Lovin, 1987).

Toutefois, il convient de considérer une dimension de l'homogénéité des individus qui n'est pas prise en compte dans notre étude. En effet, les individus suivis ont fait un choix similaire d'environnement d'études et de conditions de vie. Ils partagent un même projet de reprise d'études après un parcours professionnel, ils choisissent de vivre ce projet avec leur famille (pour certains) et intègrent à cet effet une université sélective mondialement renommée. Par conséquent, il est probable que l'homophilie exogène soit en fait plus explicative que ce nous avons pu déceler via les attributs socio-démographiques mis à notre disposition.

Le récit des relations entrelacées

A l'analyse de nos résultats, nous constatons que l'évolution des relations amicales déclarées correspond en partie seulement à l'évolution des liens sociaux multiplexes entre les individus. L'observation des échanges sociaux effectifs à distance et des co-présences, apporte un éclairage complémentaire sur l'évolution des relations amicales. Les échanges effectifs répondent à des considérations pratiques, instrumentales et à d'autres logiques que celles des déclarations d'amitié. Ces liens se concrétisent au sein de structures d'opportunités qui émergent au fil du temps lors de rencontres et de partage d'activités. Cet entrelacement des relations déclaratives et effectives peut être appréhendé par la configuration multiplexe des réseaux.

Nous avons observé que ces réseaux s'influencent mutuellement de façon significative dans leur dynamique d'évolution temporelle par des effets structuraux multiplexes, tels que :

- l'impact des choix relationnels des individus au sein d'un réseau sur leurs choix au sein d'un autre réseau (hypothèse H_s4).
- la concordance des choix relationnels d'un réseau à l'autre (hypothèse H_s5).

Toutefois, cet entrelacement des descripteurs des relations déclarées et observées est complexe à interpréter. Par conséquent, plutôt que d'évoquer des concordances ou des discordances entre les trois réseaux qui restituent les relations sociales, nous préférons qualifier cet entrelacement comme une « anastomose »[25] des récits et des espaces de sociabilité. Une dissection plus fine de cette anastomose nécessite des investigations supplémentaires. Les propositions de Battiston *et al.* (2014) concernant des mesures structurales multidimensionnelles spécifiques pour les réseaux multiplexes, développées dans le champ de la physique et appliquées aux sciences humaines, sont de possibles pistes de poursuite de nos investigations.

Etudes comparables

Les études qui portent sur les relations d'amitiés et les phénomènes d'homophilie sont nombreuses et les résultats sont bien établis en ce qui concerne les attributs de sexe, d'âge, d'origine ethnique et de niveau socio-économique (Maisonneuve et Lamy, 1993 ; McPherson *et al.*, 2001). Nous avons cité et commenté diverses études dans le cadre de notre revue bibliographique et nos résultats convergent vers ces constats. Nous soulignons, peut-être de façon plus saillante, l'intérêt de prendre en compte l'homophilie expliquée par les facteurs structuraux mais aussi la complexité du traitement multiplexe des données issues de sources d'observations variées.

Toutefois, nous pouvons évoquer cinq limites méthodologiques de notre étude, en développant plus en détail la dernière.

Tout d'abord, nous avons fait le choix de symétriser les matrices des données relationnelles afin de constituer réseaux sociaux non dirigés. Par conséquent, nous n'avons pas exploité la réciprocité des relations dans les appariements affinitaires. Plusieurs auteurs ont souligné l'importance de la réciprocité dans les échanges de ressources instrumentales et expressives (Molm *et al.*, 2007) ainsi que de son influence sur les comportements entre pairs (Almaatouq *et al.*, 2016a ; Almaatouq *et al.*, 2016b).

---

[25] Anastomose : connexion entre des structures, des organes, des espaces. Communication entre vaisseaux sanguins (16ème siècle, médical). Métaphorique : faire se réunir, se joindre par anastomose (www.cnrtl.fr/definition/anastomoser).

En outre, l'inégale intensité des liens d'amitié réciproques dans les relations inter et intra-ethniques est soulignée par Vaquera et Gao (2008) par exemple. Par ailleurs, dans l'étude des relations amicales, l'articulation de la réciprocité avec d'autres mécanismes structuraux est documentée (par exemple : Block, 2014). Cependant, comme nous l'avons justifié, compte-tenu de la base de données exploitée, ce choix nous a permis d'optimiser le volume des effectifs suivis au cours des étapes d'observation et d'éviter de réduire la perte d'information relative aux données manquantes dans nos analyses descriptives et nos modélisations longitudinales.

Deuxièmement, nous avons choisi d'appliquer des seuils sur l'intensité des liens. Ceci nous a permis de dégager la structure dorsale des réseaux (*backbone*) en éliminant des liens d'intensité de faible intensité, que nous avons considérés négligeables pour une étude sur les relations amicales. Nous avons qualifié de « bruit » l'intensité inférieure à ces seuils. Toutefois, tout choix de seuil peut être interrogé et nous évoquerons plus avant un problème méthodologique de fond sur ce qui est parfois considéré comme du « bruit » ou un « biais » dans les observations sociales.

Troisièmement, selon l'article qui décrit globalement l'expérimentation menée par le MIT (Aharony *et al.*, 2011) environ un tiers des individus suivis sont en couple avec des pairs proches (*significant other*). Toutefois, les données proposées sur la plateforme de distribution des bases de données ne contenaient pas des informations telles que les liens conjugaux, les revenus des individus ou leurs âges en détail. Ces informations sont probablement jugées stratégiques et confidentielles pour certains financeurs de l'expérimentation intéressés par les résultats de l'expérimentation sur l'influence des pairs dans l'adoption d'applications pour téléphones mobiles (Pan *et al.*, 2011a) ou de pratiques sportives régulières (Aharony *et al.*, 2011). La mise à disposition de ces informations anonymisées nous aurait très certainement permis d'affiner nos analyses.

Quatrièmement, le questionnaire sur les liens d'amitiés est basé sur des réponses à hiérarchiser selon 7 niveaux possibles : « *0: I don't know this person ; 1: I know of this person ; 2: This person is an acquaintance ; 3: This person is a friend-low ranking ; 4: friend-medium ranking ; 5. Friend-high ranking ; 6: close friend ; 7: family or as close to me as a family member* ». Compte-tenu de la diversité des origines culturelles des étudiants interrogés, il est difficile d'écarter des fluctuations dans l'interprétation du sens et des nuances de ces sept niveaux de réponses. Nous avons précédemment souligné cette difficulté dans la qualification d'une relation amicale et affinitaire (Marsden, 1988 ; Maisonneuve et Lamy, 1993 ; Van de Bunt *et al.*, 1999).

Enfin, cinquièmement, nous souhaitons développer la limite méthodologique relative à une étude multiplexe des relations sociales qui croisent des données de nature différentes.

A notre connaissance, aucun des travaux publiés à ce jour ne fait état d'une modélisation longitudinale de réseaux complets qui croise des liens d'amitié avec des données issues d'observation d'interactions synchrones à distance et de détection de proximités physiques. La comparaison de notre recherche avec d'autres travaux est par conséquent difficile sur cette dimension.

Des auteurs tels que Burt (2000), McPherson *et al.* (2001) soulignaient, il y a une vingtaine d'année, le besoin de plus d'études longitudinales croisant des données variées sur les relations multiplexes entre les individus, afin de progresser sur la compréhension de l'homophilie dans les réseaux sociaux. Les méthodes traditionnellement utilisées pour recueillir des observations du type multiplexe sont celles du questionnaire et de l'observation directe. Toutefois, depuis quelques décennies, avec le développement des plateformes numériques sur l'Internet et la diffusion massive des téléphones portables, de multiples possibilités sont offertes aux chercheurs intéressés pour modéliser des interactions sociales à partir de l'utilisation conjointe de plusieurs sources d'observations. Ainsi, il est de plus en plus fréquent que des « expérimentations naturelles » croisent des données empiriques issues de recueil d'observations qui utilisent différentes méthodes de mesure (Lazer *et al.*, 2009 ; Anahory *et al.*, 2011; Barrat *et al.*, 2010 ; Smieszek *et al.*, 2011 ; Stehlé *et al.*, 2011, Mastrandrea *et al.*, 2015 ; Starnini *et al.*, 2017). Cette utilisation conjointe de données de nature différente révèle des écarts parfois qualifiés de « biais » par des chercheurs. Ce qui peut apparaître comme des discordances dans le croisement des récits de la sociabilité des individus n'a pas, de tout évidence, les mêmes conséquences pour des sociologues (Bernard et Killworth, 1977, Bernard *et al.*, 1980, Bernard *et al.*, 1984) que pour des épidémiologistes (Smieszek *et al.*, 2011 ; Stehlé *et al.*, 2011).

Les premiers pourront parfois conclure que ces discordances sont susceptibles de « déprimer les chercheurs qui se basent sur la remémoration en substitut de collecte de données par observation participante » (Russel Bernard *et al.*, 1984, p.496) et que « les individus ne savent pas avec précision avec qui ils communiquent » (Bernard *et al.*, 1981 p.89). Au contraire, des auteurs soulignent que « les données de remémoration peuvent être utilisées pour prédire des aspects structuraux de données issues de l'observation » (Romney et Faust, 1982, p.303) mais, toutefois, ils dégagent une conclusion éloquente : *« Proof is in the pudding »* (Romney et Faust, 1982, p. 302).

Les seconds nous apparaissent actuellement les usagers les plus fréquents des possibilités offertes par les développements technologiques pour modéliser les interactions entre les individus. Ces développements technologiques permettent de plus en plus aisément de combiner des données issues de questionnaires « contacts », de traces d'interactions sur l'Internet et de capteurs de proximité. De nombreuses plateformes de recherches interdisciplinaires proposent des travaux et les données afférentes (*Reality Mining*[26], *Socio-Patterns*[27], *Socio-Metrics badges data* [28]). Par exemple, la plateforme « *Socio-Patterns* » est riche d'expérimentations menées sur les dynamiques sociales, principalement dans le champ de l'épidémiologie. Ces expérimentations naturelles mixent différentes méthodes et outils de recueil simultané de données longitudinales dans des contextes variés (écoles, hôpitaux…).

Parmi ces études, des constats répétés soulignent des congruences limitées et parfois difficiles à expliquer entre l'information donnée par les questionnaires, les relevés des co-présences détectées par des capteurs de proximité (Smieszek *et al.*, 2011 ; Smieszek *et al.*, 2014) et les interactions sur les plateformes sociales numériques (Barrat *et al.*, 2010 ; Mastrandrea *et al.*, 2015). Ce que nous choisissons de nommer des dissonances plutôt que des biais entre les interactions restituées et effectives, sont le plus souvent attribuées soit au déclin de la mémoire, soit à une utilisation sélective des médias d'interactions (Smoreda *et al.*, 2007 ; Bailliencourt *et al.*, 2007), soit à une remémoration dilective qui privilégie les liens d'amitiés forts au détriment des liens faibles (Mastrandrea *et al.*, 2015). Ces derniers auteurs soulignent, comme d'autres chercheurs (Barrat *et al.*, 2010), qu'il est bien difficile de distinguer ce qui peut relever de biais, d'écarts, de recouvrement ou de compléments dans ce que restituent les informations ainsi multiplexées.

Cependant, l'étude de Mastrandrea *et al.* (2015) ainsi que celles de Stehlé *et al.* (2011, 2013) et Fournet et Barrat (2014) qui portent sur l'observation des interactions entre des étudiants (ou élèves) avec un dispositif de recueil de données longitudinales des co-présences, mentionnent toutes un effet de la structure où se déroulent les opportunités d'interactions. Toutefois, celle-ci est encore difficile à disséquer. Ainsi, Stehlé *et al.* (2013) concluent à une homophilie de sexe dans l'observation des proximités spatiales au sein d'un très jeune public en milieu scolaire, alors que Fournet et Barrat (2014) n'en décèlent pas chez un public adolescent en Lycée où elle est pourtant connue comme plus prégnante dans cette tranche d'âge (Maisonneuve et Lamy, 1993 ; McPherson *et al.*, 2001 ; Pearsons, 2005).

[26] http://realitycommons.media.mit.edu/realitymining.html.
[27] http://www.sociopatterns.org/ (Barrat *et al.*, 2008).
[28] http://www.sociometricsolutions.com.

Toutefois, ces études n'observent pas les interactions qui se déroulent en dehors de la classe, dans l'espace public ouvert ou privé, ni sur une période suffisamment longue pour discerner des évolutions. Ceci limite de surcroît la comparaison avec notre propre recherche.

Néanmoins, des travaux récents ont été conduits afin d'apprécier la robustesse des modélisations statistiques qui utilisent simultanément des données de natures variées, notamment celles issues de la détection des co-présences (Starnini *et al.*, 2017) pour lesquelles nous avons souligné la complexité d'interprétation.

Starnini *et al.* (2017) traitent les données issues de différentes méthodes et outils utilisés pour détecter les interactions de proximité entre des individus (capteurs « *Bluetooth* », badges à fréquences radiophoniques et transmissions infrarouges). L'objectif est d'apprécier si les caractéristiques statistiques de base et structurales des réseaux répondent effectivement à une loi de distribution robuste et commune ou si des divergences apparaissent selon le contexte et l'instrument d'observation. Les auteurs concluent que, globalement, les différentes bases de données relationnelles ont des propriétés statistiques similaires en ce qui concerne, d'une part, la restitution des comportements des individus au regard de la distribution de la durée des co-présences et, d'autre part, la distribution des écarts temporels entre deux co-présences détectées pour un individu donné.

Toutefois, Starnini *et al.* (2017) soulignent également deux points moins satisfaisants qui nous interpellent concernant les limites méthodologiques de notre propre étude. D'une part, la distribution du temps cumulé des co-présences par un individu et, d'autre part, la corrélation entre le temps cumulé d'interaction pour un individu et le nombre d'interactions de l'individu (le degré du nœud) « présentent une évolution assez bruitée » (Starnini *et al.*, 2017, p.13). Les études comparées par les auteurs utilisent des capteurs de co-présences basés sur des capteurs « *Bluetooth* » (portée de détection inférieure à 10 mètres), des badges à fréquences radiophoniques (portée inférieure à 1,5 mètres) et des émetteurs/capteurs à infrarouges (cône de détection de 1 mètre de hauteur). Or, de façon souvent non nuancée les auteurs des expérimentations qui exploitent les données issues de ces capteurs font référence à une détection d'interaction dyadique en « *face-à-face* », plutôt qu'à une détection de proximité (ou de co-présence) temporelle de deux ou plusieurs individus dans un espace restreint. C'est le cas pour les onze travaux qui exploitent la base de données que nous avons utilisée, ainsi que pour plusieurs des travaux cités précédemment. Cela constitue, selon nous, un problème de fond méthodologique que nous avons évoqué lors de nos propres analyses.

En effet, à l'exception des données issues des capteurs infra-rouges, qui détecte effectivement une présence entre deux individus en ligne directe - dans un cône de tolérance plus ou moins étroit (Starnini *et al.*, 2017, p. 5) - les autres capteurs fournissent des données de co-présences qui ne peuvent pas avec une certitude satisfaisante être qualifiées d'observations « *d'interaction en face-à-face* », compte-tenu du périmètre de portée de ces capteurs.

En outre, à l'exception des badges sociométriques (Lepri *et al.*, 2012) qui intègrent un microphone, aucun de ces capteurs ne peut garantir qu'il y a eu une communication verbale ni même non verbale, c'est-à-dire une interaction sociale effective entre les individus dont la co-présence en un même lieu est détectée en même temps.

Ainsi, dans notre étude, la détection des proximités entre les individus est réalisée par un capteur « *Bluetooth* » intégré dans les téléphones portables des individus dont les interactions sociales sont observées. Par ailleurs, les concepteurs de l'expérimentation mentionnent que « le signal pouvant traverser les murs des appartement voisins et détecter une co-présence, les données collectées entre minuit et sept heures du matin ne sont pas retenues dans la base de données » (Aharony *et al.*, 2011, p.651). Par conséquent, nous relevons une limite méthodologique de fond dans l'interprétation sociologique de ce que restitue ce type de données relationnelles. Nous considérons que ces données agrègent tout à la fois des informations sur des interactions effectives entre les individus et des informations qui concernent deux ou plusieurs individus qui partagent en même temps un espace restreint sans forcément interagir entre eux. Dans ce dernier cas, soit ils sont à proximité (logement proche, cheminement proche « face-à-dos »), soit ils participent - activement ou passivement - à une même activité (spectacle, conférence, sport) sans interaction inter-personnelle. Compte-tenu des méthodes et des outils utilisés pour les collectes de données de proximité spatiale entre les individus, nous faisons l'hypothèse que ces données relationnelles « quantifient » plus probablement un effet global relatif aux structures d'opportunités et aux foyers d'activités (Blau, 1977 ; Feld, 1981), qu'un effet lié à une interaction dyadique confirmée.

Les enjeux de la granulosité et de l'hétérogénéité des données

Il n'est pas mis en doute que l'évolution majeure de ce qui est parfois nommée la « numérisation » du monde (*datatification)* a provoqué et continue de provoquer des changements méthodologiques significatifs dans le traitement « traditionnel » des données (Damon, 2013 ; Mayer-Schönberger et Cukier, 2014). Ces évolutions ont été accélérées par les avancées technologiques qui ont « boosté » la transformation des objets en données.

Ces évolutions méthodologiques ont augmenté la granulosité de l'exploration scientifique en rendant accessible chaque grain élémentaire, c'est-à-dire le contenu atomique des parties d'un objet : les lettres qui composent les mots des pages d'un livre, le nombre de secondes d'un face-à-face entre deux individus. Le même phénomène s'observe dans des champs toujours plus variés : la numérisation des images, du son, des émotions, des interactions et des comportements humains et *in fine* de l'ensemble de l'activité humaine observable (*quantified self mouvement*). Un changement de paradigme provient ainsi de l'augmentation de la puissance du microscope d'observation des objets, qu'ils soient inertes, vivants ou sociaux. Ce changement entraine des évolutions méthodologiques dans l'appréhension des nouvelles particules élémentaires ainsi accessibles. Il appartient, selon nous, au « scientifique des données » en sciences sociales par ses connaissances en statistiques, en sociologie et sa déontologie, d'apprécier la significativité des données qui sont exploitées et de différencier, dans la mesure du possible, ce qui relève du bruit ou du signe (le « *pattern* » au sens statistique) en sociométrie. C'est-à-dire de révéler le sens des données dans les formes utilisées pour leur collecte : « *faire parler les données* » (Cardon, 2012. p.139) en intégrant les nuances plus ou moins importantes induites par les méthodes acquisition.

Le sens de la règle centenaire des seuils de significativité statistique (Fisher, 1925) est un guide robuste pour aider à identifier la valeur d'une observation sociale quantifiée. Ce guide permet de distinguer ce qui est informatif de ce qui est du bruit. Néanmoins, les nouvelles possibilités offertes à la sociométrie pour le recueil des interactions sociales, virtuelles ou physiques, conduisent à une abondance massive de données qui peuvent sembler similaires mais qui, lorsque l'on examine au scalpel les méthodes de recueil, apparaissent suffisamment hétérogènes pour interroger leur agrégation en un même contenu informatif indistinctement.

En l'état des méthodes, des outils et de la nature hétérogène des données mises à disposition, il nous apparaît bien difficile de déplier dans nos résultats ce qui peut relever, d'une part, des choix des acteurs rationnels motivés par la recherche de ressources sociales, des préférences individuelles expliquées par une homophilie d'attributs acquis et hérités et, d'autre part, des conséquences de l'organisation sociale de la structure qui catalyse ou contraint les opportunités d'interactions.

Afin de contribuer à mieux éclairer la dynamique des interactions sociales par l'identification d'effets propres d'une homophilie socio-démographique, d'une homophilie induite et d'une homophilie structurale endogène au réseau, il serait utile d'harmoniser *plus encore* les méthodes, les outils de mesure ainsi que les vocables utilisés, afin de désigner de façon homogène ce qui relève d'une interaction dyadique lors de la quantification d'une observation sociale.

Bibliographie

Almaatouq, A., Radaelli, L., Pentland, A., Shmueli, E. (2016a). Are you your friends' friend? Poor perception of friendship ties limits the ability to promote behavioral change. PloS one, 11(3), e0151588.

Almaatouq, A., Radaelli, L., Pentland, A., & Shmueli, E. (2016b). The Role of Reciprocity and Directionality of Friendship Ties in Promoting Behavioral Change. In International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (pp. 33-41). Springer, Cham.

Abiteboul, S. (2012). Sciences des données : de la logique du premier ordre à la Toile. Leçon inaugurale du cours au Collège de France. [En ligne] URL : https://www.college-de-france.fr/site/serge-abiteboul/inaugural-lecture-2012-03-08-18h00.htm.

Adler, P., Kwon, S.W. (2002). Social Capital: Prospects for a New Concept. *The Academy of Management Review*, 27(1), 17-40.

Aharony, N., Pan, W., Ip, C., Khayal, I., Pentland, A. (2011). Social FMRI: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, 7(6), 643-659.

Allan, G. (1979). A sociology of friendship and kinship. G. Allen & Unwin.

Allan, G. (1998). Friendship, Sociology and Social Structure. *Journal of Social and Personal Relationships*, 15(5), 685–702.

Alter, N. (dir.). (2006). Sociologie du monde du travail. Paris : Presses Universitaires de France.

Altshuler, Y., Aharony, N., Pentland, A., Elovici, Y., Cebrian, M. (2011a). Stealing Reality: When Criminals Become Data Scientists (or Vice Versa). *IEEE Intelligent Systems*, 26(6), 22-30.

Altshuler, Y., Aharony, N., Fire, M., Elovici, Y., Pentland, A. (2011b). *Incremental learning with accuracy prediction of social and individual properties from mobile-phone data*. [En ligne] URL : https: //arxiv.org/abs/1111.4645.

Altshuler, Y., Fire, M., Aharony, N., Elovici, Y., Pentland, A. (2012). How Many Makes a Crowd? On the Correlation between Groups' Size and the Accuracy of Modeling. In *Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction* (pp. 43-52). [En ligne] URL : https://www.media.mit.edu/publications/how-many-makes-a-crowd-on-the-correlation-between-groups-size-and-the-accuracy-of-modeling/.

Amati, V., Meggiolaro, S., Rivellini, G., Zaccarin, S. (2018). Social relations and life satisfaction: the role of friends. *Genus*, 74(1), 7.

Axelrod, R.-M. (1984). The evolution of cooperation. New York: Basic Books.

Bailliencourt (De), T., Beauvisage, T., Smoreda, Z. (2007). La communication interpersonnelle face à la multiplication des technologies de contact. Réseaux, (6), 81-115.

Barrat, A., Cattuto, C., Szomszor, M., Van den Broeck, W., Alani, H. (2010). Social dynamics in conferences: analyses of data from the Live Social Semantics application. *In International semantic web conference* (pp. 17-33). Springer, Berlin, Heidelberg.

Barnes-Mauthe, M., Arita, S. (2012). A Network Analysis of Fisher's Social Capital and the Effects of Ethnic Diversity in Hawaii's Longline Fishery. *Pelagic Fisheries Research Program, SOEST, 12-03.*

Barnes, J. (2013). Classes sociales et réseaux dans une île de Norvège. Réseaux, 182(6), 209-237. doi:10.3917/res.182.0209.

Battiston, F., Nicosia, V., Latora, V. (2014). Structural measures for multiplex networks. *Physical Review E*, 89(3), 032804.

Berge, C. (1958). Théorie des graphes et ses applications Dunod.

Berger, P., Luckmann, T. (1966). The social construction of reality. Londres, Penguin.

Bernard, R. H, Killworth P. (1977). Informant accuracy in social network data II. *Human Communication Research*, 4 (1), 3-18.

Bernard, R. H., Killworth P., Sailer L. (1979/1980). Informant accuracy in social network data IV : a comparison of clique-level structure in behavorial and cognitive network data. *Social Networks*, 2, 191-218.

Bernard, R. H., Killworth P., Kronenfeld D., Sailer L. (1984). The problem of accuracy : the validity of retrospective data. *Ann. Rev. Anthropol*., 13, 495-517.

Blase, J.(1986). Socialization as humanization: One side of becoming a teacher. Sociology of Education, 59, 100-1.

Bes, M-P. (2016). Des alumni à l'université française : comment créer ex-nihilo un réseau de diplômés ? *Recherches anthropologiques et sociologiques*, 47 (1), 89-110.

Bes, M-P. (2009a). Un réseau Ecole d'ingénieurs-industrie : l'intérêt du multi-niveaux. Congrès *Etudes sur les réseaux sociaux à Toulouse*, Janvier 2009, Toulouse, France.

Bes, M-P. (2009b). Entretenir son réseau : le cas des ingénieurs. *3ème Congrès de l'Association Française de Sociologie*, Avril 2009, Paris, France.

Bidart, C., Degenne, A., Grossetti, M. (2011). *La vie en réseau : Dynamique des relations sociales*. Paris : Presses Universitaires de France.

Bidart, C. (1997). L'amitié, un lien social. Paris: La Découverte.

Blau, P.M. (1977). A Macrosociological Theory of Social Structure. *American Journal of Sociology*, 83 (1), 26-54.

Blau, P.M. (1993). Multilevel structural analysis. *Social Networks*, 15(2), 201-215.

Blau, P. M. (1994). Structural contexts of opportunities. University of Chicago Press.

Block, P., Grund, T. (2014). Multidimensional Homophily in Friendship Networks. *Network science*, 2(2), 189–212.

Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks*, 40, 163-173.

Boissevain, J. (1974). Friends of friend s: Networks, manipulators and coalitions. St. Martin's Press.

Boser, B.E., Guyon, I.M., Vladimir, N., Vapnik, A. (1992). Training Algorithm for Optimal Margin Classifier. In *Fifth Annual Workshop on Computational Learning Theory* (pp.144-152). Pittsburgh: ACM.

Bourdieu, P. (1980). Le capital social. *Actes de la recherche en sciences sociales*, 31, 2-3.

Brain, R. (1976). Friends and lovers. Grafton Books.

Burt, R.S. (1992). Structural Holes : The Social Structure of Competition. Harvard University Press, Cambridge.

Burt, R.S. (1995). Le capital social, les trous structuraux et l'entrepreneur. *Revue française de sociologie*, 36(4), 599-628.

Burt, R.S. (2000). The Network structure of social capital. *Research in Organizational Behavior*, 22, 345-423.

Burt, R.S. (2001a). Structural holes versus network closure as social capital. In Lin, N., Cook, K. S., Burt, R.S, *Social capital : theory and research* (pp.1-30). New York : Aldine de Gruyter.

Burt, R.S. (2001b). The social capital of structural holes. In Guillén, M. F, Collins, R., England, P., Meyer, M. (eds), *New directions in Economic sociology. Developments in an emerging field* (pp.201-250). New York : Russel Sage Fondation.

Burt, R.S. (2019). Nan Lin and social capital. In Burt, R.S., Bian, Y., Song, L., Lin, N. *Social capital, social support and stratification : an analysis of the sociology of Nan Lin* (pp. 1-72). Forthcoming 2020. Edward Elgar Publishing.

Cardon, D., et Smoreda, Z. (2014). Réseaux et les mutations de la sociabilité. Réseaux, (2), 161-185.
Cardon, D. (2015). A quoi rêvent les algorithmes. Nos vies à l'heure: Nos vies à l'heure des big data. Le Seuil.

Cattuto C., Van den Broeck W., Barrat A., Colizza V., Pinton J.F., Vespignani A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS one*, 5(7), e11596.

Centola, D., González-Avella, J. C., Eguíluz, V. M., San Miguel, M. (2007). Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups. *Journal of Conflict Resolution*, 51(6), 905–929.

Clauset, A., Newman, M. E. J., Moore, C. (2004). Finding Community Structure in Very Large Networks. *Physical Review*, E70, 066111.

Coleman, J.S. (1988). Social capital in the creation of human capital. *American journal of sociology, 94*, 95-120.

Coleman, J. S. (1994). Foundations of social theory. Harvard university press.

Crozier, M., Friedberg, E. (1977). *L'acteur et le système : les contraintes de l'action collective*. Paris : Seuil.

Davis, A., Gardner, B. B., Gardner, M. R. (2009). Deep South: A social anthropological study of caste and class. Univ of South Carolina Press.

Degenne, A., et Forsé, M. (2004). Les réseaux sociaux. *Paris : Armand Colin.*

Degenne, A. (2013). L'analyse des réseaux sociaux. Un survol à travers quelques jalons. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 118(1), 22-43.

Deschenau, F., Laflamme, C. (2009). Réseau social et capital social : une distinction conceptuelle nécessaire illustrée à l'aide d'une enquête sur l'insertion professionnelle de jeunes Québécois. *SociologieS.* [En ligne] URL : http://journals.openedition.org/sociologies/2902.

DiPrete, T.A, Gelman, A., McCormick, T., Teitler, J., Zheng, T. (2011). Segregation in Social Networks Based on Acquaintanceship and Trust. *American Journal of Sociology*, 116 (4), 1234-83.

Dussuc, B., Geindre, S. (2012). Capital social, théorie des réseaux sociaux et recherche en PME : une revue de la littérature. *11ème congrès CIFEPME (Congrès International francophone en entrepreneuriat et PME)*, Oct 2012, Brest, France. [En ligne] URL : https://halshs.archives-ouvertes.fr/halshs-00747912/document.

Durlauf, S. N. (1999). The case "against" social capital. Social Systems Research Institute, University of Wisconsin.

Eagle N., Macy M., Claxton R. (2010). Network diversity and economic developpment. *Science*, 328(5981), 1029-1031.

Eagle, N., Pentland, A.S., Lazer, D. (2009). Inferring Friendship Network Structure by Using Mobile Phone Data. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), 15274-8).

Ebbesen, E. B., Kjos, G. L., Konečni, V. J. (1976). Spatial ecology: Its effects on the choice of friends and enemies. *Journal of Experimental Social Psychology*, 12(6), 505-518.

Favereau, O., Lazega, E. (eds). (2002). *Conventions and structures in economic organization : markets, networks and hierarchies. New horizons in institutional and evolutionary economics series*. Cheltenham, UK, Northampton, MA, USA : Edward Elgar.

Federico de la Rúa (de), A. (2004). L'analyse longitudinale de réseaux sociaux totaux avec Siena—Méthode, discussion et application. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 84(1), 5-39.

Feld, S.L. (1981). The focused organization of social ties. *American Journal of Sociology*, 86(5), 1015-1035.

Feld, S.L. (1982). Social structural determinant of similarity among associates. *American Sociological Review*, 47(6), 797-801.

Ferrand, A. (1993). Un modèle interprétatif : « l'amitié-âge ». L'analyse des réseaux personnels. Habilitation à Diriger des Recherches. Université Lille 1.

Festinger, L., Schachter, S., Back, K. (1950). The spatial ecology of group formation. In L. Festinger,

S. Schachter, & K. Back (Eds.), Social pressure in informal groups (pp.33-59). Stanford, CA:

Stanford University Press.

Field, J. (2008). *Social Capital*. New York: Routledge.

Fischer, C. S., Jackson, R., Stueve, C. A., Gerson, K., Jones, L. M. (1977). Networks and places. Free Press.

Fischer, C.S. (1982). What do we mean by 'friend'? an inductive study Social Networks. *Social Networks,* 3(4), 287-306.

Flap, H. (2002). No man is an island : the research programme of a social capital theory. In Favereau, O., et Lazega, E. (eds), *Conventions and structures in economic organization : markets, networks and hierarchies. New horizons in institutional and evolutionary economics series* (pp. 29- 59). Cheltenham, UK, Northampton, MA, USA : Edward Elgar.

Forsé, M. (2001). Rôle spécifique et croissance du capital social. Revue de l'OFCE, 76(1), 189-216.

Forsé, M., Parodi, M. (2004). La priorité du juste, éléments pour une sociologie des choix moraux, Paris , PUF.

Forsé, M. (2008). Définir et analyser les réseaux sociaux : Les enjeux de l'analyse structurale. *Informations sociales*, 147(3), 10-19.

Fournet, J., Barrat, A. (2014). Contact patterns among high school students. *PloS one,* 9(9), e107878.

Friedkin, N. (1993). Structural Bases of Interpersonal Influence in Groups : A Longitudinal Case Study. *American Sociological Review*, 58(6), 861-872.

Génois, M., et Barrat, A. (2018). Can co-location be used as a proxy for face-to-face contacts ? *EPJ Data Science*, 7(11). [En ligne] URL : https://hal.archives-ouvertes.fr/hal-01788667.

Godechot, O., Mariot, N. (2004). Les deux formes du capital social : structure relationnelle des jurys de thèses et recrutement en science politique. *Revue française de sociologie*, 45(2), 243-282.

Granovetter, M. S. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, 78(6) ,1360-1380.

Granovetter, M. S. (1983). The Strength of Weak Ties : a network theory revisited. *Sociological Theory*, 1, 201-233.

Granovetter, M. S. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 91, 481-510.

Grossetti, M., Bes, M-P., (2003). Dynamiques des réseaux et des cercles. Encastrements et découplages. *Revue d'économie industrielle*, 103, 43-58.

Grossetti, M. (2015). Note sur la notion d'encastrement. *SociologieS*. [En ligne] URL : http://journals.openedition.org/sociologies/4997.

Heider, F. (1946). Attitudes and Cognitive Organization. *Journal of Psychology*, 21, 107-12.

Héran, F. (1988). La sociabilité, une pratique culturelle. Economie et statistique, 216, 3-22.

Hobbes, T. (1651) Leviathan or the Matter, Forme and Power of A Commonwealth Ecclesiastical and civil. London. Printed for Andrew Crooke.

Holland, P. W., Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative group studies*, 2(2), 107-124.

Holland, P. W., Leinhardt, S. (1976). Local structure in social networks. *Sociological methodology*, 7, 1-45.

Hurlbert, J., Haines, V., Beggs, J. (2000). Core Networks and Tie Activation: What Kinds of Routine Networks Allocate Resources in Nonroutine Situations? *American Sociological Review*, 65, 598-618.

Ihaka, R., Gentleman, R. (1996). « R : A language for data analysis and graphics », Journal of Computational and Graphical Statistics, vol. 5, 299-314.

Jaccard, P. (1900). Contributions au problème de l'immigration post-glaciaire de la flore alpine. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37, 547–579.

Killworth P., Russel Bernard H. (1979/1980). Informant accuracy in social network data III : a comparison of triadic structure in behavorial and cognitive data. *Social Networks* (2), 10-46.

Killworth P., Mc Carty C., Russel Bernard H., House M. (2006). The accuracy of small world chains in social networks. *Social networks* (28), 85-96.

Kossinets, G., Watts, D.J. (2009). Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*, 115(2), 405-450.

Lai, G., Lin, N., Leung, S.-Y. (1998). Network resources, contact resources, and status attainment. *Social Networks*, 20(2), 159-178.

Lazarsfeld, P. F., Merton, R. K. (1954). Friendship as a social process : A substantive and methodological analysis. Freedom and control in modern society, 18(1), 18-66.

Lazega, E. (1998). *Réseaux sociaux et structures relationnelles*. Paris : Presses Universitaires de France.

Lepri, B., Staiano, J., Rigato, G., Kalimeri, K., Finnerty, A., Pianesi, F., Sebe, N., Pentland, A. (2012). The sociometric badges corpus: A multilevel behavioral dataset for social behavior in complex organizations. In 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing (pp. 623-628). IEEE.

Lévesque M., White, D. (1999). Le concept de capital social et ses usages. *Lien social et politiques*, 41, 23-33.

Lévesque, M., White, D. (2001). Capital social, capital humain et sortie de l'aide sociale pour des prestataires de longue durée. *The Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 26(2), 167-192.

Lin, N. (1982) Social Resources and Instrumental Action. In Marsden P. V., et Lin, N. (eds). *Social Structure and Network Analysis* (pp. 131-45). Beverly Hills, CA : Sage.

Lin, N. (1995). Les ressources sociales : une théorie du capital social. *Revue française de sociologie*, 36(4), 685-704.

Lin, N. (1999a). Building a theory of social capital. *Connections*, 22(1), 28-51.

Lin, N. (1999b) Social Networks and Status Attainment. *Annual Review of Sociology*, 25, 467-487.

Lin, N., Fu, Y., Hsung, R.M. (2001) The position generator: Measurement techniques for investigations of social capital. In Lin, N., Cook, K. and Burt, R.S. (eds), *Capital : Theory and Research* (pp. 57-81). New York :Walter de Gruyter.

Lin, N. (2005). A Network Theory of Social Capital. Handbook on Social Capital, edited by Dario Castiglione, Jan van Deth and Guglielmo Wolleb.

Lin, N. (2008). A Network Theory of Social Capital. In Castiglione, D., van Deth, J., et Wolleb, G. (eds), *The handbook of social capital* (pp. 1- 25). Oxford : Oxford University Press.

Lospinoso J., Snijders, T. (2015). Model selection and RSiena. Script, version 13. [En ligne] URL : http://www.stats.ox.ac.uk/~snijders/siena/siena_news.htm.

Lospinoso, J. A., Schweinberger, M., Snijders, T. A., Ripley, R. M. (2011). Assessing and accounting for time heterogeneity in stochastic actor oriented models. Advances in data analysis and classification, 5(2), 147–176. [En ligne] URL : https://doi.org/10.1007/s11634-010-0076-1.

Maillochon, F. (2012). 9 – Pourquoi l'analyse des réseaux ?. Dans : Serge Paugam éd., L'enquête sociologique (pp. 187-206). Paris cedex 14, France: Presses Universitaires de France. doi:10.3917/puf.paug.2012.01.0187.

Maisonneuve, J., Lamy, L. (1993). Psychosociologie de l'amitié. FeniXX.

Marsden, P. V. (1981). Models and methods for characterizing the structural parameters of groups. *Social Networks*, 3(1), 1-27.

Marsden, P.V., Lin, N. (1982). *Social Structure and Network Analysis*. Beverly Hills, CA : Sage.

Marsden, P.V. (1988). Homogeneity in confiding relations. *Social Networks*, 10(1), 57-76.

Mastrandrea, R., Fournet, J., Barrat, A. (2015). Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys. *PloS one*. [En ligne] URL : https://doi.org/10.1371/journal.pone.0136497

Mauss, M. (1923). Essai sur le don forme et raison de l'échange dans les sociétés archaïques. L'Année sociologique (1896/1897-1924/1925), 1, 30-186.

Mayer-Schöberger, V., et Cukier, K. (2014). Big data. La révolution des données est en marche. Paris : Robert Laffont.

McPherson, J. M., Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American sociological review*, 370-379.

McPherson, M., Smith-Lovin, L., Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, 415-444.

Membiela-Pollán, M., Pena-López, J.-A. (2017). Clarifying the concept of social capital through its three perspectives: individualistic, communitarian and macro-social. *European Journal of Government and Economics*, 6(2), 146-170.

Mercklé, P. (2011). La sociologie des réseaux sociaux. Paris : La Découverte.

Merton, R. K. (1947). The social psychology of housing. In W. Dennis (ed.) (pp. 163-217). Current Trends in Social Psychology. Pittsburgh: Univ. of Pittsburgh Press.

Mollica, K.A., Gray, B., Treviño, L.K (2003). Racial homophily and its persistence in newcomers' social networks. *Organization Science*, 14, 123-136.

Molm, L. D., Schaefer, D. R., Collett, J. L. (2007). The value of reciprocity. *Social Psychology Quarterly*, 70(2), 199-217.

Montjoye (de) Y.A., Quoidbach J., Robic, F., Pentland A. (2013) Predicting Personality Using Novel Mobile Phone-Based Metrics. In Greenberg A.M., Kennedy W.G., Bos N.D. (eds), *Social Computing, Behavioral-Cultural Modeling and Prediction. Computer Science*, 7812.

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology*, 107(3), 679-716.

Moreno, J. L. (1934). Nervous and mental disease monograph series, no 58. Who shall survive ?: A new approach to the problem of human interrelations. Nervous and Mental Disease Publishing Co. https://doi.org/10.1037/10648-000.

Moturu, S.T., Khayal, I., Aharony, N., Pan, W., Pentland, A. (2011). Using social sensing to understand the links between sleep, mood, and sociability. *Proc. IEEE International Conference on Social Computing*. Cambridge, MA. [En ligne] URL : http://hd.media.mit.edu/tech-reports/TR-670.pdf.

Newman M.E.J (2004). Fast algorithm for detecting community structure in networks. *Physical Review*, E69, 066133.

Nishikata, T., Hardjono T., Pentland A. (2018). Social Capital Accounting. [En ligne] URL : https://www.media.mit.edu/publications/social-capital-accounting.

Oloritun, R. O, Madan, A., Pentland, A., Khayal, I. (2013). Identifying Close Friendships in a Sensed Social Network. *Procedia - Social and Behavioral Sciences*, 79, 18-26.

Pan, W., Aharony, N., Pentland, A. (2011a). Composite social network for predicting mobile apps installation. *Computer Science, Physics*. [En ligne] URL : https://arxiv.org/abs/1106.0359.

Pan, W., Aharony, N., Pentland, A. (2011b). Fortune monitor or fortune teller : Understanding the connection between interaction patterns and financial status. *IEEE Third International Conference on Privacy, Security, Risk and Trust.* [En ligne] URL : http://hd.media.mit.edu/tech-reports/TR-667.pdf.

Paulhus, D. L., Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), Handbook of research methods in personality psychology (p. 224–239). The Guilford Press.

Pearson, M., Steglich, C.E., Snijders, T.A. (2006). Homophily and assimilation among sport active adolescent substance users. *Connection Science*, 27, 47-63.

Pena-López, J.-A., Sánchez-Santos, J.-M. (2017). Individual social capital: Accessibility and mobilization of resources embedded in social networks. *Social Networks*, 49, 1-11.

Ponthieux, S. (2006). Le capital social. Paris : La Découverte.

Pons P., Latapy M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications (JGAA)*, 10(2), 191-218.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual review of sociology*, 24(1), 1-24.

Putnam, R. (1993). The prosperous community: Social capital and public life. *The american prospect, 13(Spring)*, Vol. 4. [En ligne] URL : http://www. prospect. org/print/vol/13.

Putnam, R. (1995). Bowling Alone. America ’s Declining Social capital. *The Journal of Democracy*, 6(1), 65-78.

Putnam, R.D., Rey, O. (2008). *Bowling en solo : le déclin du capital social américain*. Conférence n°27, 417-440. [En ligne] URL : https://halshs.archives-ouvertes.fr/halshs-00943612

Read, J.M., Eames, K.T., Edmunds, W.J. (2008). Dynamic social networks and the implications for the spread of infectious disease. *Journal of The Royal Society Interface*, 5(26), 1001-1007.

Reichardt J., Bornholdt S. (2006). Statistical mechanics of community detection. *Physical Review*, E74, 016110.

Ripley, R.M, Snijders, T.A.B, Boda, Z., Vörös, A., Preciado, P. (2019, 2020). Manual for RSeina. [En ligne] URL : http://www.stats.ox.ac.uk/~snijders/siena/.

Romney K. A., Faust K. (1984). Predicting the structure of a communications network from recalled data. *Social Networks*, 4, 285-304.

Scott, J. (1991). Social network analysis a Handbook (2nd éd.). London: Sage Publications.

Smieszek, T., Burri, E. U., Scherzinger, R., Scholz, R. W. (2012). Collecting close-contact social mixing data with contact diaries: reporting errors and biases. Epidemiology & infection, 140(4), 744-752.

Smieszek, T., Barclay, V. C., Seeni, I., Rainey, J. J., Gao, H., Uzicanin, A., Salathé, M. (2014). How should social mixing be measured: comparing web-based survey and sensor-based methods. BMC infectious diseases, 14(1), 136.

Smoreda, Z., Beauvisage, T., de Bailliencourt, T., Assadi, H. (2007). Saisir les pratiques numériques dans leur globalité. Réseaux, 145-146(6), 19-43.

Snijders, T. (1999). Prologue to the Measurement of Social Capital. *La Revue Tocqueville/The Tocqueville Review*, 20(1), 27-44.

Snijders, T. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, 31, 361-395.

Snijders, T. A. B., van de Bunt, G. G., Steglich, C. (2010). Introduction to actor based models for network dynamics. *Social Networks*, 32, 44-60.

Snijders, T. A. B., Lomi, A., Torlo, V.J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, 35, 265-276.

Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., Pentland, A. Friends don't lie : Inferring Personality Traits from Network Structure. In *UbiComp '12 Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp.321-330). New York : ACM New York.

Starnini, M., Lepri, B., Baronchelli, A., Barrat, A., Cattuto, C., Pastor-Satorras, R. (2017). Robust modeling of human contact networks across different scales and proximity-sensing techniques. In *International Conference on Social Informatics* (pp. 536-551). Springer, Cham.

Stauder, J. (2014). Friendship networks and the social structure of opportunities for contact and interaction. *Social science research*, 48, 234-250.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J-F., Quaggiotto, M., Broeck (van den), W., Regis, C., Lina, B., Vanhems, P. (2011). High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. *PloS one*, 6(8), e23176.

Stehlé, J., Charbonnier, F., Picard, T., Cattuto, C., Barrat, A. (2013). Gender homophily from spatial behavior in a primary school: A sociometric study. *Social Networks*, 35(4), 604-613.

Urberg, K. A., Degirmencioglu, S. M., Tolson, J. M. (1998). Adolescent Friendship Selection and Termination: The Role of Similarity. *Journal of Social and Personal Relationships*, 15(5), 703–710.

Vaquera, E., Kao, G. (2008). Do You Like Me as Much as I Like You? Friendship Reciprocity and Its Effects on School Outcomes among Adolescents. *Social science research*, 37(1), 55-72.

Vandenberghe, F. (2009). III. Sociologie et épistémologie. In F. Vandenberghe (éd.), La sociologie de Georg Simmel (pp. 35-50). Paris : La Découverte.

Van De Bunt, G.G., Van Duijn, M.A. Snijders, T.A. (1999). Friendship Networks Through Time: An Actor-Oriented Dynamic Statistical Network Model. *Computational & Mathematical Organization Theory*, 5, 167–192. https://doi.org/10.1023/A:1009683123448.

Van Duijn, M. A., Zeggelink, E. P., Huisman, M., Stokman, F. N., Wasseur, F. W. (2003). Evolution of sociology freshmen into a friendship network. *Journal of Mathematical Sociology*, 27(2-3), 153-191.

Verbrugge, L.-M. (1977). The structure of adult friendship choices. *Social Forces*, 56, 576-597.

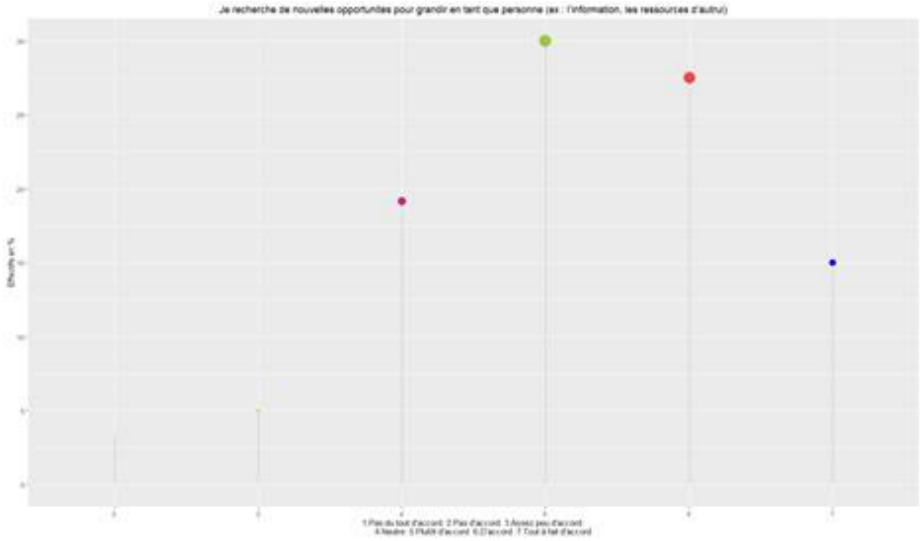
Verbrugge, L. (1979). Multiplexity in Adult Friendships. *Social Forces*, 57(4), 1286-1309.

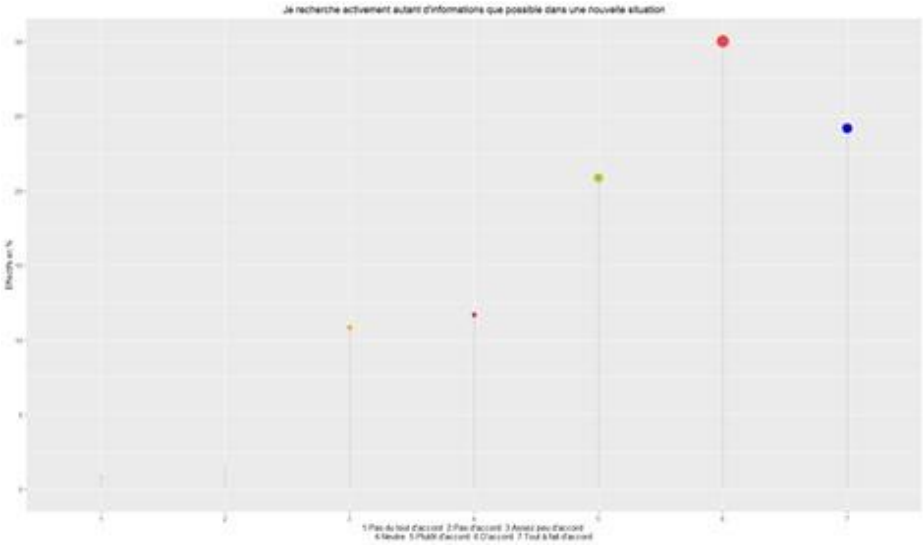
Wrong, D. (1961). The over-socialized conception of man in modern sociology. *American sociological review* (26), 183-193.

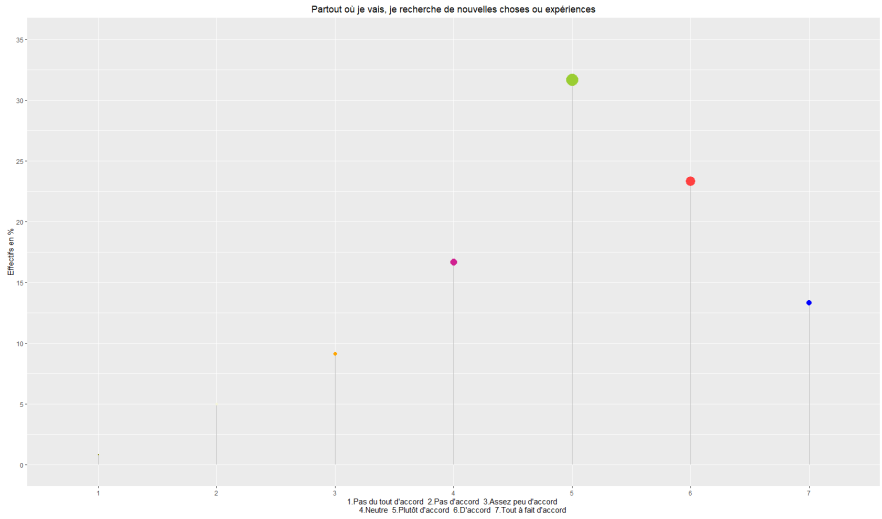
7. Annexes

A.1 Personnalité des étudiants suivis

Graphique A.1 : distribution des réponses sur la personnalité des étudiants (effectif : 117 individus)





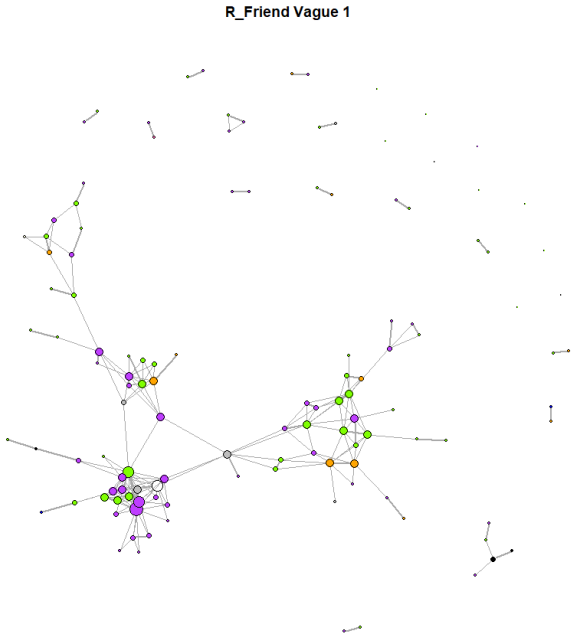


A.2 Métriques du réseau « R_Friend », avec un seuil sur l'intensité des liens égal à 3

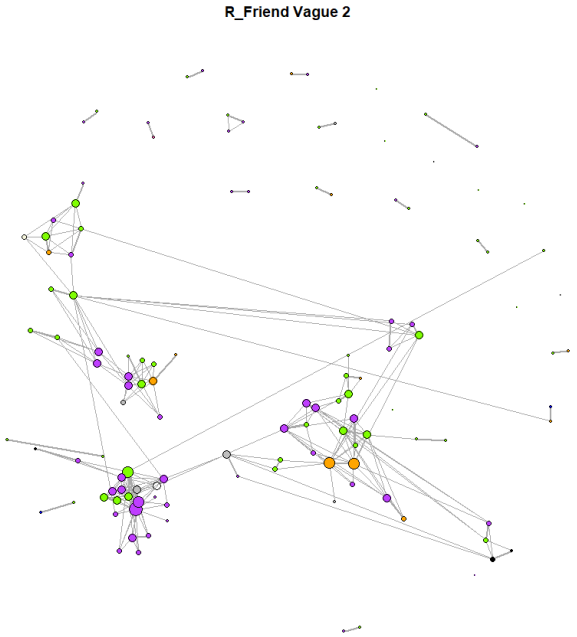
Tableau n°A.1 : principales métriques du réseau « R_Friend », avec un seuil sur l'intensité des liens égal 3.

Réseau « R_Friend » non dirigé seuil : intensité des liens ≥3	**Vague n°1**	**Vague n°2**	**Vague n°3**	**Vague n°4**
Taille du réseau	123 nœuds, 490 liens	123 nœuds, 571 liens	123 nœuds, 574 liens	123 nœuds, 653 liens
Densité	0,065	0,076	0,076	0,087
Longueur moyenne du chemin le plus court *(Average Path Length)*	2,834	2,787	2,743	2,591
Coefficient global d'agglomération *(Global Clustering Coefficient)*	0,428	0,424	0,415	0,394
Coefficient local moyen d'agglomération *(Local Clustering Coefficient)*	0,535	0,522	0,533	0,524
Centralisation de Freeman normalisée	0,188	0,186	0,202	0,232
Centralité de degré normalisée				
Moyenne	0,065	0,076	0,076	0,087
Ecart-type	0,058	0,064	0,064	0,070
Degré maximal	0,254	0,26	0,278	0,319
Centralité d'intermédiarité normalisée - valeur maximale *(standardized betweenness centrality)*	0,208	0,200	0,130	0,10
Origine ethnique des 10 premiers étudiants « *gatekeepers* »	Blanche (5), Hispanique (1), Asiatique (3), Noire (1)	Blanche (7), Hispanique (1), Asiatique (1), Moyen-Orientale (1)	Blanche (4), Asiatique (6)	Blanche (4), Asiatique (4), Moyen-Orientale (1) inconnue (1)

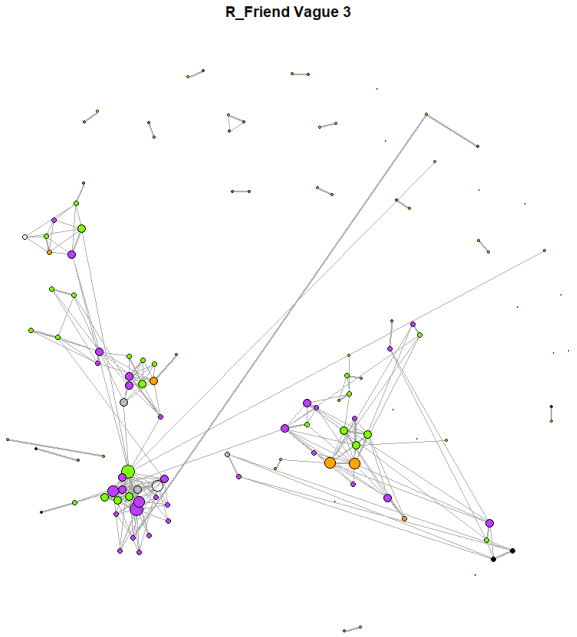
Graphique n°A.2 : représentation de l'évolution du réseau non dirigé « R_Friend » pour les 4 vagues, seuil sur l'intensité des liens d'amitié de 3. L'épaisseur des liens est proportionnelle à l'intensité de la relation et la taille des nœuds est proportionnelle aux degrés.



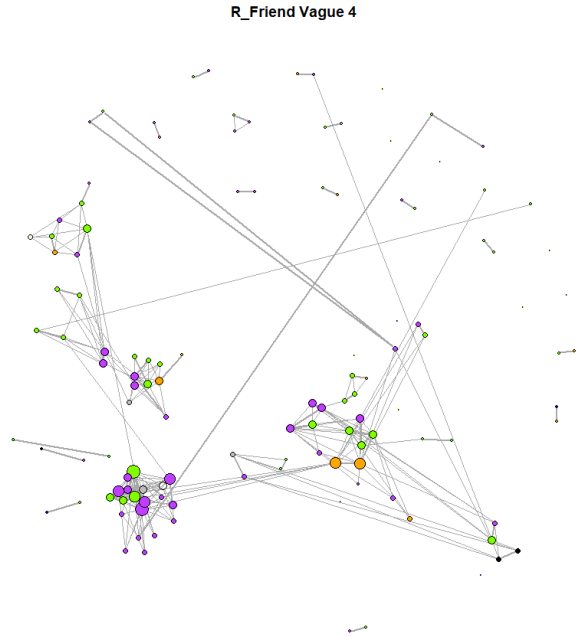
Vague n°1 : 123 nœuds, 208 liens d'intensité 5 à 7



Vague n°2 : 123 nœuds, 243 liens d'intensité 5 à 7



Vague n°3 : 123 nœuds, 254 liens d'intensité 5 à 7



Vague n°4 : 123 nœuds, 255 liens d'intensité 5 à 7

Origine ethnique : Asiatique (violet), Noire (noir), Hispanique (orange), Israélienne (beige), Moyen-Orientale (gris), Métissée (bleu), Amérindienne (rose), Blanche (vert).

Tableau n°A.2 : métriques réseau non orienté « R_Call », sans seuil.

Réseau « R_Call » non dirigé	**Vague n°1**	**Vague n°2**	**Vague n°3**
Taille	112 nœuds, 251 liens	112 nœuds, 259 liens	112 nœuds, 250 liens
Densité	0,040	0,042	0,040
Average Path Length	2,93	3,03	2,95
Clusters - nombre [taille]	4 [106, 2, 2, 2]	2 [110, 2]	3 [108, 2, 2]
Cohésion des réseaux			
Transitivité globale *Global Clustering Coefficient*	0,21	0,23	0,20
Transitivité locale moyenne *Local Clustering coefficient*	0,51	0,50	0,50
Centralisation de Freeman normalisée	0,40	0,36	0,38
Centralité de degré normalisée			
Moyenne	0,04	0,04	0,04
Ecart-type	0,05	0,05	0,05
Degré maximal	0,01	0,01	0,01
Centralité d'intermédiarité normalisée valeur maximale	0,56	0,69	0,61

Graphique n°A.3 : distribution des degrés des nœuds du réseau « R_Call », sans seuil.
(effectif : 112 individus)

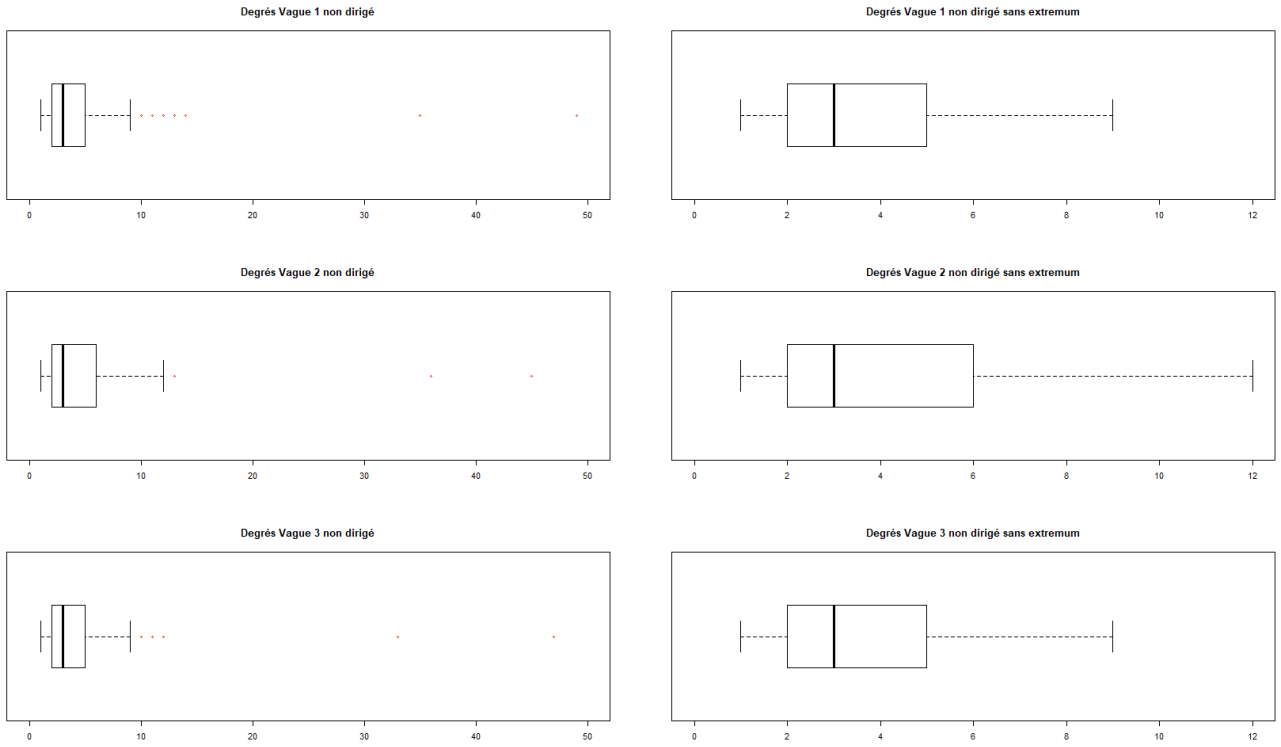


Tableau n°A.3 : distribution des contacts téléphoniques pour le réseau non orienté « R_Call », sans seuil.

Vague n°1			Vague n°2			Vague n°3		
[nombre de jours]	**Nombre de liens**	**%**	**[nombre de jours]**	**Nombre de liens**	**%**	**[nombre de jours]**	**Nombre de liens**	**%**
[1 à 2)	50,00	19,90	[1 à 2)	63,00	24,30	[1 à 2)	52	20,80
[2 à 5)	64,00	25,50	[2 à 5)	59,00	22,80	[2 à 5)	62	24,80
[5 à 18)	74,00	29,50	[5 à 18)	71,00	27,40	[5 à 18)	70	28,00
[18 à 83]	63,00	25,10	[18 à 84]	66,00	25,50	[18 à 81]	66	26,40
Total	251	100%		259	100%		250	100%

Lecture : pour la vague n°1, en moyenne 25,10% des individus se sont contactés au moins un fois par téléphone sur une période de18 à 83 jours.

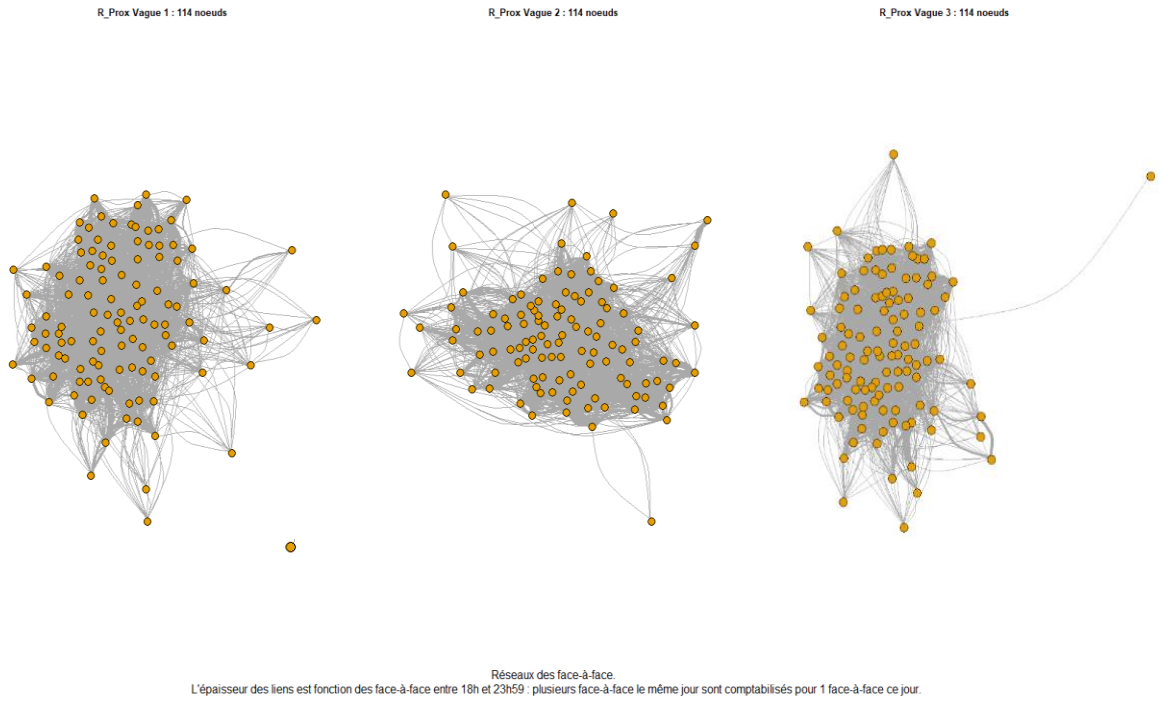
Tableau n°A.4 : métriques du réseau non orienté « R_Prox, sans seuil.

Réseau « R_Prox » non orienté	**Vague n°1**	**Vague n°2**	**Vague n°3**
Taille du réseau	114 nœuds, 2859 liens	114 nœuds, 3049 liens	114 nœuds, 3482 liens
Densité	0,44	0,47	0,54
Average Path Length	1,56	1,52	1,47
Transitivité globale *Global Clustering Coefficient*	0,58	0,64	0,68
Transitivité locale moyenne *Local Clustering coefficient*	0,60	0,66	0,70
Centralisation de Freeman normalisée	0,34	0,28	0,35
Moyenne	0,44	0,47	0,54
Ecart-type	0,16	0,19	0,19
Degré maximal	0,79	0,75	0,89
Centralité d'intermédiarité normalisée valeur maximale	0,03	0,03	0,02

Tableau n°A.5 : nombre d'heures de co-présences par vague d'observation. Réseau « R_Prox », sans seuil. (Effectif : 114 individus).

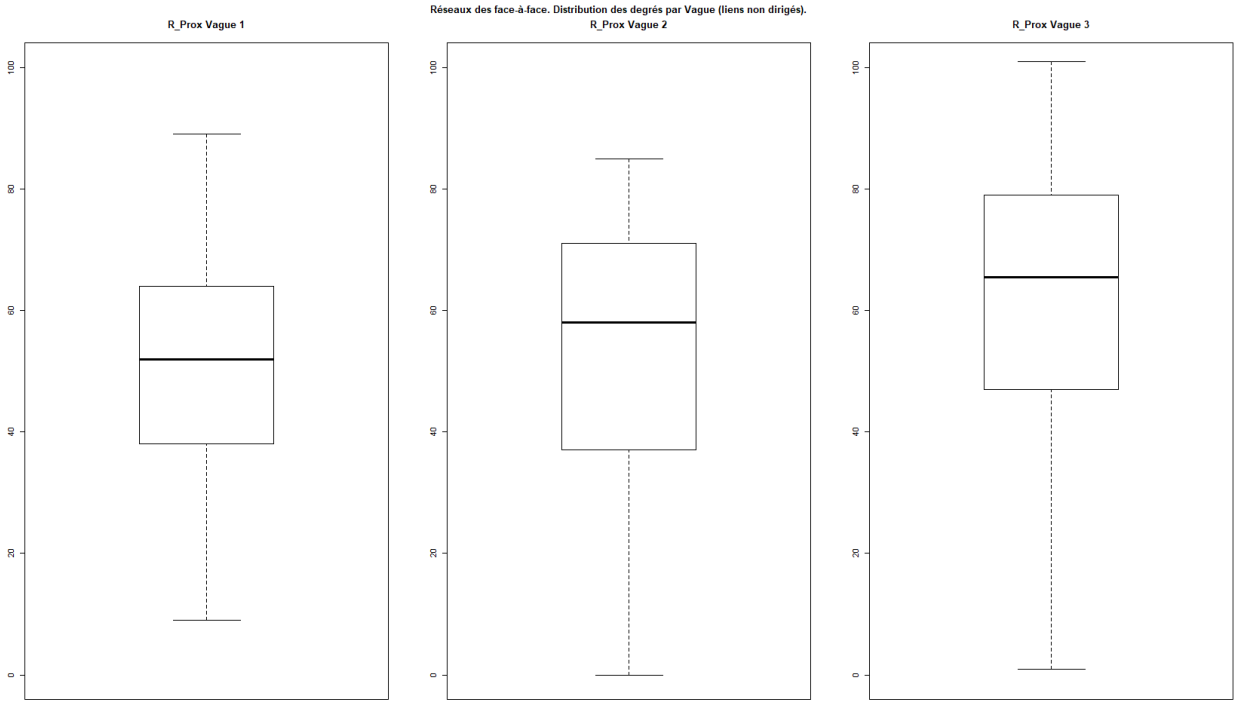
	Min.	**1 Qu.**	**Médiane**	**3 Qu.**	**Max.**	**Moyenne**
Vague n°1	0,080	0,080	0,330	1,830	218,920	7,342
Vague n°2	0,080	0,17	0,58	2,42	421,92	11,79
Vague n°3	0,080	0,17	0,42	1,83	380	10,64

Graphique n°A.4 : représentation de l'évolution du réseau « R_Prox » pour les 3 vagues (sans seuil).

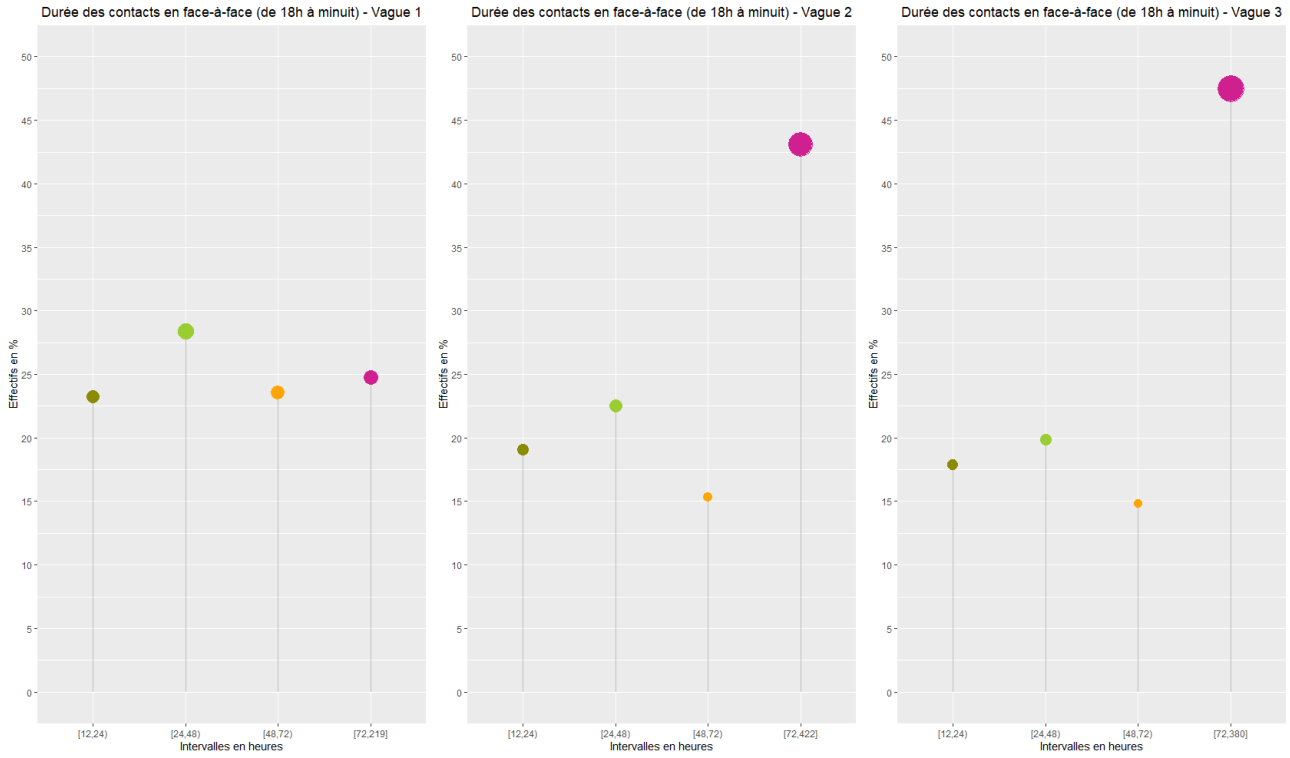


Vague n°1 : 114 nœuds, 2859 liens non dirigés	Vague n° 2 : 114 nœuds, 3049 liens non dirigés	Vague n° 3 : 114 nœuds, 3482 liens non dirigés
Densité : 0,443	Densité : 0,47	Densité : 0,54
Average path length : 1,56	Average path length : 1,52	Average path length : 1,47

Graphique n°A.5 : distribution des degrés des nœuds du réseau non orienté « R_Prox », sans seuil. (effectif : 114 individus).



Graphique n°A.6 : durée des contacts du réseau « R_Prox » par intervalles d'heures (seuil : ≥12h). (effectif : 114 individus)



Postulats pour le modèle SAOM (*Stochastic Actor-Oriented Models for network dynamics*)

Les postulats sur l'action temporelle des acteurs (I) impliqués dans la structure relationnelle (II) sont résumés ci-dessous, à partir des travaux de Snijders *et al.* (2001, 2010) et de la synthèse de Federico de la Rúa (2004).

I. Concernant les acteurs, il est fait l'hypothèse de base que chaque acteur poursuit ses propres objectifs sous la contrainte générée par son entourage constitué par les autres acteurs. Les postulats sont les suivants :

   a. tous les acteurs ont la même probabilité de réaliser un changement dans leurs liens
   b. chaque acteur maitrise le choix de ses relations. L'acteur réalise des changements dans l'établissement de ses liens en fonction des caractéristiques des autres acteurs, sa position dans le réseau et sa perception de l'ensemble de la structure relationnelle.
   c. à chaque moment, l'acteur a la possibilité de faire un seul changement. Ainsi une seule relation peut changer à chaque moment et le processus se décompose en composants élémentaires ce qui autorise une modélisation « opérationnalisable » (Holland et Leinhardt, 1977).
   d. les acteurs ne se coordonnent pas préalablement à leurs choix : ces choix sont réalisés de façon séquentielle, sans négociation ni concertation préalable. Il est ainsi supposé que la réciprocité d'un lien n'est pas instantanée : un des acteurs de la dyade relationnelle prend l'initiative unilatérale à un instant t_i de créer un lien, et ce lien pourra devenir ultérieurement réciproque à l'initiative de l'autre acteur de la dyade au temps t_{i+n}.
   e. les acteurs forment des choix selon un stratégie de maximisation à court terme de l'utilité des liens établis, maintenus ou rompus
   f. les acteurs du réseau peuvent évaluer leur position dans le réseau et tenter d'acquérir une configuration relationnelle qu'ils considèrent positive
   g. les acteurs ont une connaissance totale des agencements des liens dans le réseau au temps t_i.
   h. les acteurs n'ont pas de mémoire du passé ou de capacité d'apprentissage

II. Concernant la structure, il est considéré que l'amplitude des effets qui influencent l'évolution du réseau ne varie pas entre deux mesures successives, si toutefois l'écart temporel n'est pas trop grand au regard des phénomènes étudiés. Les postulats sont alors énoncés comme suit :

   i. des changements latents et continus ont lieu de façon non observée entre les temps t_i et t_{i+1}.
   j. les changements ont lieu l'un après l'autre.
   k. les états observé du réseau à un temps t_{i+1} sont une fonction seulement de l'état observé au temps t_i : ainsi, il est considéré que « le passé pertinent est contenu dans le présent » (de la Rúa, 2004 , p.7)
   l. l'évolution assure une dépendance dans les changements effectués par les acteurs.
   m. l'évolution du réseau est conditionnée par la première observation de celui-ci. Cette première observation n'est pas modélisée : elle est utilisée comme point de départ et conditionne la procédure d'estimation statistique. Par conséquent, seuls les changements entre 2 observations sont modélisés et non les déterminants des changements qui ont conduits à la première observation. Cette procédure est proposée par Snijders (2001) afin de dépasser les limites des méthodes d'estimation classique pour ce type de modèles complexes.

La modélisation inclue un résidu qui exprime la part due aux phénomènes aléatoires et non expliqués par les éléments modélisés, qui s'écartent des postulats posés. Ce résidu, plus ou moins important, explique l'écart entre la configuration structurale empirique observée et la configuration théorique calculée.

Modèles SIENA

Modélisation longitudinale du réseau des échanges téléphoniques « R_Call »

Liens sans seuil sur l'intensité des liens.

Tableau A.6 : indicateurs de densité Réseau « R_Call », sans seuil.

Vague d'observation	**n°1**	**n°2**	**n°3**
Densité	0,040	0,042	0,040
Degré moyen	4,482	4,625	4,464
Nombre de liens	251	259	250

Tableau A.7 : modélisation longitudinale du réseau « R_Call », sans seuil.

Réseau « R_Call » sans seuil	Forcing & initiative models		Pairwise forcing model	
Effets	paramètre	(é.t)	paramètre	(é.t)
taux 1	1,102	(0,119)	0,121	(0,012)
taux 2	1,128	(0,121)	0,123	(0,013)
Effets structuraux				
degré (densité)	–1,933***	(0,146)	–3,505***	(0,287)
appariement degré[1] équivalent				
activité + popularité				
triades transitives				
4-cycles (1)				
Effets des co-variables				
sexe Alter	–0,137	(0,231)	–0,421	(0,412)
même sexe	–0,113	(0,146)	–0,151	(0,273)
ethnie A Alter	–0,243	(0,262)	–0,843†	(0,475)
même ethnie A	0,179	(0,152)	0,217	(0,287)
ethnie B Alter	–0,404	(0,248)	–1,068*	(0,459)
même ethnie B	0,131	(0,154)	0,258	(0,275)
nombre d'enfants d'Alter	0,187	(0,139)	0,198	(0,260)
similarité du nombre d'enfants	0,356	(0,227)	0,550	(0,412)
inter. sexe * nbre d'enfants	0,340	(0,251)	0,497	(0,437)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1
ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001;
(é.t) : écart-type.
(1) racine carrée des degrés.

Tableau A.8 : modélisation longitudinale du réseau « R_Call », sans seuil.

Réseau « R_Call » sans seuil	Forcing & initiative models		Forcing & initiative models	
Effets	paramètre	(é.t)	paramètre	(é.t)
taux 1	1,239	(0,141)	1,242	(0,142)
taux 2	1,281	(0,140)	1,284	(0,146)
Effets structuraux				
degré (densité)	–1,461***	(0,266)	–1,619***	(0,294)
appariement degré[1] équivalent	–0,370***	(0,064)	–0,335***	(0,067)
activité + popularité	0,102***	(0,018)	0,094***	(0,019)
triades transitives	0,844***	(0,088)	1,005***	(0,138)
4-cycles (1)			–0,030	(0,021)
Effets des co-variables				
sexe Alter	0,104	(0,221)	0,101	(0,239)
même sexe	–0,140	(0,145)	–0,128	(0,154)
ethnie A Alter	0,400	(0,297)	0,377	(0,337)
même ethnie A	0,203	(0,159)	0,202	(0,162)
ethnie B Alter	0,316	(0,299)	0,295	(0,338)
même ethnie B	0,074	(0,165)	0,063	(0,162)
nombre d'enfants d'Alter	0,369*	(0,144)	0,336*	(0,149)
similarité du nombre d'enfants	0,407†	(0,216)	0,398†	(0,226)
inter. sexe * nbre d'enfants	0,403	(0,274)	0,381	(0,274)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1
ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001
(é.t) : écart-type.
(1)racine carrée des degrés.

Modélisation longitudinale du réseau des co-présences « R_Prox »

Liens sans seuil sur l'intensité des liens.

Tableau A.9 : indicateurs de densité Réseau « R_Prox », sans seuil.

Vague d'observation	**n°1**	**n°2**	**n°3**
Densité	0,444	0,473	0,541
Degré moyen	50,158	53,491	61,088
Nombre de liens	2859	3049	3482

Tableau A.10 : modélisation longitudinale du réseau « R_Prox », sans seuil.

Réseau « R_Prox » sans seuil	Forcing & initiative models		Forcing & initiative models		Forcing & initiative models	
Effets	paramètre	(é.t)	paramètre	(é.t)	paramètre	(é.t)
taux 1	32,156	(1,063)	43,245	(2,162)	42,176	(1,958)
taux 2	28,254	(0,869)	38,318	(1,951)	43,512	(2,460)
Effets structuraux						
degré (densité)	0,049*	(0,025)	–2,433***	(0,082)	–0,906***	(0,122)
appariement degré[1] équivalent			0,031***	(0,001)		
activité+popularité					–0,003*	(0,001)
triades transitives					0,046***	(0,003)
Effets des co-variables						
sexe Alter	–0,038	(0,038)	–0,001	(0,033)	–0,001	(0,031)
même sexe	–0,014	(0,025)	–0,013	(0,025)	–0,011	(0,024)
ethnie A Alter	0,265***	(0,056)	0,086†	(0,047)	0,090†	(0,048)
même ethnie A	0,007	(0,032)	–0,003	(0,030)	–0,001	(0,028)
ethnie B Alter	0,351***	(0,052)	0,095*	(0,046)	0,109*	(0,045)
même ethnie B	0,044	(0,031)	0,031	(0,028)	0,036	(0,028)
nombre d'enfants d'Alter	–0,089**	(0,030)	0,019	(0,025)	0,005	(0,027)
similarité du nombre d'enfants	0,002	(0,042)	–0,003	(0,038)	–0,006	(0,040)
inter. sexe * nbre d'enfants	–0,045	(0,049)	–0,004	(0,042	–0,002	(0,042)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001; (é.t) : écart-type.
(1) : racine carrée des degrés.

Tableau A.11 : modélisation longitudinale du réseau « R_Prox », sans seuil.

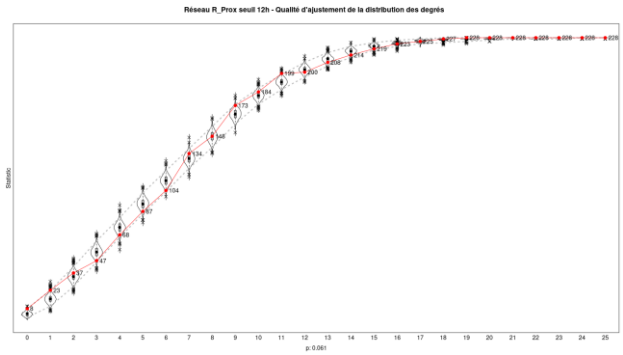
Réseau « R_Prox » sans seuil	Pairwise forcing model		Pairwise forcing model		Pairwise forcing model	
Effets	paramètre	(é.t)	paramètre	(é.t)	paramètre	(é.t)
taux 1	0,570	(0,018)	0,909	(0,046)	0,919	(0,045)
taux 2	0,502	(0,016)	0,870	(0,043)	1,068	(0,061)
Effets structuraux						
degré (densité)	0,087†	(0,050)	–4,990***	(0,202)	–1,579***	(0,314)
appariement degré[1] équivalent			0,062***	(0,002)		
activité+popularité					–0,008*	(0,003)
triades transitives					0,099***	(0,007)
Effets des co-variables						
sexe Alter	–0,079	(0,069)	–0,018	(0,066)	–0,022	(0,061)
même sexe	–0,028	(0,051)	–0,029	(0,047)	–0,023	(0,048)
ethnie A Alter	0,535***	(0,108)	0,183†	(0,109)	0,191†	(0,107)
même ethnie A	0,014	(0,060)	–0,007	(0,062)	–0,008	(0,059)
ethnie B Alter	0,708***	(0,116)	0,204*	(0,094)	0,223*	(0,104)
même ethnie B	0,093	(0,062)	0,064	(0,062)	0,075	(0,053)
nombre d'enfants d'Alter	–0,184**	(0,060)	0,032	(0,055)	0,012	(0,051)
similarité du nombre d'enfants	0,011	(0,084)	–0,010	(0,085)	–0,014	(0,077)
inter. sexe * nbre d'enfants	–0,088	(0,104)	–0,014	(0,090)	–0,023	(0,089)

Pour tous les modèles présentés : convergence des « *t ratios* » < 0,1 et ratio maximum de convergence < 0,25.
Seuils de significativité statistique : † p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001; (é.t) : écart-type.
(1) racine carrée des degrés.

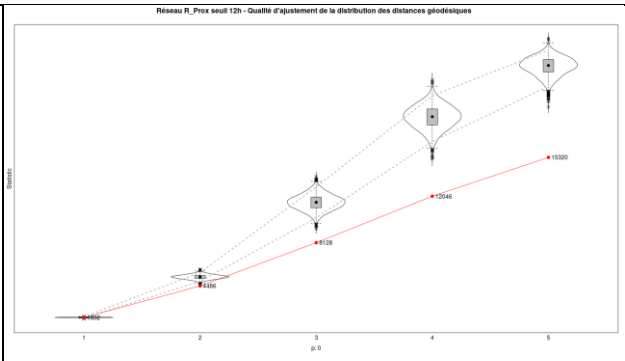
La figure A.7 présente les résultats de qualité d'ajustement du modèle P3 ci-dessous :

R_Prox (seuil : ≥ 12h de co-présence /vague) ~ Degrés +Σ Popularité & Activité + Triades transitives
+ co-variables d'Alter (sexe, ethnie, nbre d'enfant)
+ égalité (sexe, ethnie) + similarité (nbre d'enfants) + interaction sexe * nbre d'enfants.

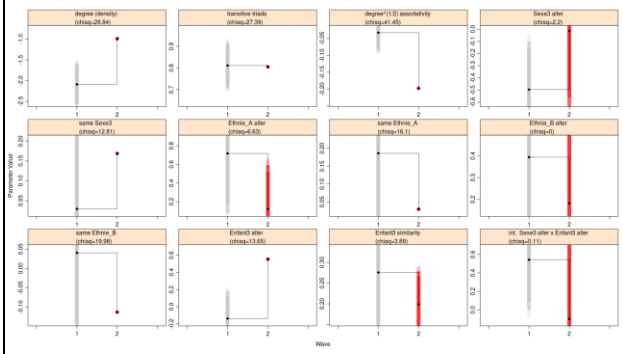
Figure n°A.7 : résultats des tests de la qualité d'ajustement du modèle P3



degrés. p-value = 0,061



distances. p-value = 0



test d'hétérogénéité temporelle

Modélisation longitudinale du réseau « R_Friend ∩ R_Call ∩ R_Prox », avec seuil sur l'intensité des liens.

Tableau A.12 : modélisation longitudinale du réseau « R_Friend∩R_Call_R_Prox », avec seuil. (à droite : FCP1, à gauche : FCP2)

Effect	par.	(s.e.)
constant R_Friend rate (period 1)	2.123	(0.203)
constant R_Friend rate (period 2)	1.303	(0.136)
R_Friend: degree (density)	−2.388 ***	(0.167)
R_Friend: Sexe3 alter	0.113	(0.174)
R_Friend: same Sexe3	−0.002	(0.130)
R_Friend: Ethnie_A alter	1.163 ***	(0.303)
R_Friend: same Ethnie_A	0.120	(0.162)
R_Friend: Ethnie_B alter	0.994 ***	(0.278)
R_Friend: same Ethnie_B	0.070	(0.146)
R_Friend: Enfant3 alter	0.573 ***	(0.132)
R_Friend: Enfant3 similarity	0.215	(0.198)
int. R_Friend: Sexe3 alter x Enfant3 alter	−0.442 †	(0.243)
R_Friend: from R_Call agreement	−0.087	(0.143)
R_Friend: R_Call to agreement	1.043 ***	(0.229)
R_Friend: R_Call	1.279 **	(0.416)
R_Friend: from R_Prox agreement	−0.098	(0.071)
R_Friend: R_Prox to agreement	0.403 ***	(0.093)
R_Friend: R_Prox	1.202 ***	(0.348)
constant R_Call rate (period 1)	0.992	(0.143)
constant R_Call rate (period 2)	0.924	(0.135)
R_Call: degree (density)	−3.107 ***	(0.242)
R_Call: Sexe3 alter	−0.612 *	(0.310)
R_Call: same Sexe3	−0.369 †	(0.201)
R_Call: Ethnie_A alter	−0.112	(0.432)
R_Call: same Ethnie_A	0.126	(0.207)
R_Call: Ethnie_B alter	−0.114	(0.412)
R_Call: same Ethnie_B	0.004	(0.215)
R_Call: Enfant3 alter	−0.255	(0.251)
R_Call: Enfant3 similarity	−0.083	(0.305)
int. R_Call: Sexe3 alter x Enfant3 alter	−0.421	(0.435)
R_Call: from R_Friend agreement	0.003	(0.038)
R_Call: R_Friend to agreement	0.083	(0.095)
R_Call: R_Friend	1.718 ***	(0.326)
R_Call: from R_Prox agreement	0.009	(0.070)
R_Call: R_Prox to agreement	0.367 **	(0.119)
R_Call: R_Prox	0.742 †	(0.379)
constant R_Prox rate (period 1)	1.370	(0.157)
constant R_Prox rate (period 2)	1.214	(0.139)
R_Prox: degree (density)	−2.053 ***	(0.184)
R_Prox: Sexe3 alter	0.089	(0.228)
R_Prox: same Sexe3	−0.214	(0.146)
R_Prox: Ethnie_A alter	0.188	(0.300)
R_Prox: same Ethnie_A	0.157	(0.174)
R_Prox: Ethnie_B alter	0.288	(0.290)
R_Prox: same Ethnie_B	0.055	(0.173)
R_Prox: Enfant3 alter	0.097	(0.167)
R_Prox: Enfant3 similarity	−0.061	(0.224)
int. R_Prox: Sexe3 alter x Enfant3 alter	0.637 *	(0.277)
R_Prox: from R_Friend agreement	−0.119 **	(0.043)
R_Prox: R_Friend to agreement	0.032	(0.145)
R_Prox: R_Friend	0.897 **	(0.310)
R_Prox: from R_Call agreement	−0.317 †	(0.164)
R_Prox: R_Call to agreement	1.613 ***	(0.389)
R_Prox: R_Call	1.626 ***	(0.485)

† p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001;
all convergence t ratios < 0.09.
Overall maximum convergence ratio 0.23.

Effect	par.	(s.e.)
constant R_Friend rate (period 1)	2.178	(0.210)
constant R_Friend rate (period 2)	1.326	(0.137)
R_Friend: degree (density)	−2.188 ***	(0.261)
R_Friend: transitive triads	0.270 ***	(0.061)
R_Friend: degree act+pop	0.039 †	(0.020)
R_Friend: degree^(1/2) assortativity	−0.108 *	(0.050)
R_Friend: Sexe3 alter	0.162	(0.161)
R_Friend: same Sexe3	−0.005	(0.120)
R_Friend: Ethnie_A alter	1.008 ***	(0.298)
R_Friend: same Ethnie_A	0.094	(0.152)
R_Friend: Ethnie_B alter	0.860 **	(0.286)
R_Friend: same Ethnie_B	0.076	(0.153)
R_Friend: Enfant3 alter	0.569 ***	(0.143)
R_Friend: Enfant3 similarity	0.241	(0.200)
int. R_Friend: Sexe3 alter x Enfant3 alter	−0.451 †	(0.233)
R_Friend: from R_Call agreement	−0.050	(0.127)
R_Friend: R_Call to agreement	0.518 **	(0.198)
R_Friend: R_Call	0.984 *	(0.425)
R_Friend: from R_Prox agreement	−0.106	(0.078)
R_Friend: R_Prox to agreement	0.257 **	(0.092)
R_Friend: R_Prox	1.397 ***	(0.390)
constant R_Call rate (period 1)	1.066	(0.154)
constant R_Call rate (period 2)	0.957	(0.150)
R_Call: degree (density)	−3.779 ***	(0.471)
R_Call: transitive triads	0.955 ***	(0.188)
R_Call: degree act+pop	0.137 ***	(0.032)
R_Call: degree^(1/2) assortativity	−0.255 **	(0.091)
R_Call: Sexe3 alter	−0.393	(0.371)
R_Call: same Sexe3	−0.452 *	(0.207)
R_Call: Ethnie_A alter	0.083	(0.511)
R_Call: same Ethnie_A	0.132	(0.229)
R_Call: Ethnie_B alter	0.078	(0.470)
R_Call: same Ethnie_B	0.055	(0.232)
R_Call: Enfant3 alter	−0.242	(0.291)
R_Call: Enfant3 similarity	−0.183	(0.317)
int. R_Call: Sexe3 alter x Enfant3 alter	0.130	(0.464)
R_Call: from R_Friend agreement	0.054	(0.044)
R_Call: R_Friend to agreement	−0.481 **	(0.169)
R_Call: R_Friend	2.387 ***	(0.425)
R_Call: from R_Prox agreement	−0.003	(0.080)
R_Call: R_Prox to agreement	0.094	(0.137)
R_Call: R_Prox	1.115 *	(0.462)
constant R_Prox rate (period 1)	1.886	(0.258)
constant R_Prox rate (period 2)	1.545	(0.203)
R_Prox: degree (density)	−0.970 **	(0.319)
R_Prox: transitive triads	0.815 ***	(0.062)
R_Prox: degree act+pop	0.019	(0.065)
R_Prox: degree^(1/2) assortativity	−0.234	(0.145)
R_Prox: Sexe3 alter	0.112	(0.223)
R_Prox: same Sexe3	−0.247	(0.150)
R_Prox: Ethnie_A alter	0.011	(0.296)
R_Prox: same Ethnie_A	0.186	(0.169)
R_Prox: Ethnie_B alter	−0.039	(0.288)
R_Prox: same Ethnie_B	−0.007	(0.174)
R_Prox: Enfant3 alter	0.005	(0.176)
R_Prox: Enfant3 similarity	−0.267	(0.233)
int. R_Prox: Sexe3 alter x Enfant3 alter	0.817 **	(0.287)
R_Prox: from R_Friend agreement	−0.090 *	(0.044)
R_Prox: R_Friend to agreement	−0.201	(0.148)
R_Prox: R_Friend	0.914 **	(0.336)
R_Prox: from R_Call agreement	−0.173	(0.166)
R_Prox: R_Call to agreement	1.092 **	(0.341)
R_Prox: R_Call	0.989 *	(0.449)

† p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001;
all convergence t ratios < 0.12.
Overall maximum convergence ratio 0.26.

Table des matières

Sommaire **3**

1. Introduction **4**

2. Revue bibliographique **10**

3. Matériau et méthode **43**

3.1. Formulation des hypothèses et proposition d'opérationnalisation 43

3.2. Présentation de l'expérimentation « Friends and Family » 46

3.2.1. Présentation de la base de données « Friends and Family » 46

3.2.2. Synthèse des études publiées 48

3.3. Positionnement et originalité de la présente étude 57

3.3.1. Construction des bases de données relationnelles 57

3.4. Méthodologie pour l'étude descriptive et longitudinale des réseaux 61

3.4.1. Le réseau des relations amicales « R_Friend » 62

3.4.2. Le réseau des contacts téléphoniques « R_Call » 66

3.4.3. Le réseau des co-présences « R_Prox » 67

3.4.4. Les réseaux multiplexes d'ordre 2 et 3 68

3.5. Modélisation longitudinale de la dynamique des réseaux 69

3.5.1. Le modèle SAOM 70

3.5.2. Spécifications des modèles SIENA mobilisés 73

3.5.3. Construction des modèles 75

3.5.4. Stratégie de modélisation 82

4. Résultats **85**

4.1. Analyses descriptives des réseaux uniplexes 88

4.1.1. Le réseau « R_Friend » 88

4.1.2. Le réseau « R_Call » 97

4.1.3. Le réseau « R_Prox » 102

4.2. Modélisation longitudinale des réseaux uniplexes et multiplexes 108

4.2.1. Le réseau « R_Friend » 109

4.2.2. Le réseau « R_Call » 114

4.2.3. Le réseau « R_Prox » 117

4.2.4. Le réseau « R_Friend ∩ R_Call » 123

4.2.5. Le réseau « R_Friend ∩ R_Prox » 130

4.2.6. Le réseau « R_Friend ∩ R_Call ∩ R_Prox » 137

5. Discussion et conclusions **140**

6. Bibliographie **152**

7. Annexes **160**