UFR Sciences Fondamentales et Biomédicales, Université Paris Cité
Mémoire de recherche dans le cadre du Master 2 Cybersécurité

Modélisation statistique de réseaux sociaux multiples

Réalisé par Quentin Guardia, quentin.guardia@etu.u-paris.fr

Sous la supervision de la tutrice du Ministère des Armées
et d'Ahmed Mehaoua, professeur à l'Université Paris Cité.

Septembre 2022

Note spéciale

Conformément à la réglementation en vigueur, certaines activités du Ministère des Armées sont sujettes à des contraintes de confidentialité. Dans cette optique, les missions de stage et les informations relatives à l'organisme d'accueil ne seront pas développées dans ce rapport. En revanche, un mémoire de recherche a été réalisé en respectant les thèmes abordés durant le stage et la formation.

Toute personne employée au sein d'un organisme relevant de la Défense nationale est tenue de respecter un devoir de réserve. Une clause de confidentialité s'applique de surcroît à tout document provenant d'une organisation militaire. La Direction au sein de laquelle j'ai effectué mon stage, n'y fait pas exception.

Par conséquent, la classification de mes activités au cours de ce stage impose de prendre des précautions quant à la présentation des tâches effectuées.
Le présent rapport se borne à donner une description générale des activités, sans entrer dans le détail des informations manipulées.

Résumé

Dans ce mémoire, un modèle statistique est proposé pour étudier un ensemble de réseaux sociaux, où interagit le même groupe d'acteurs. On appelle ces réseaux intriqués des « réseaux multiples ». Au-delà d'étudier les liens au sein de chacun des réseaux, le modèle proposé extrapole les informations entre les réseaux afin d'obtenir de meilleurs résultats en termes de prédiction et de qualité d'ajustement. Le modèle s'appuie sur une extension hiérarchique du modèle d'espace latent, en supposant l'existence d'une position sociale « globale » pour chaque acteur, ce qui permet d'avoir différentes approches pour chaque rôle social. La proposition a été mise en pratique avec plusieurs jeux de données réels, en considérant divers types de relation.

Mots clés: Statistique Bayésienne, Réseaux, Chaîne de Markov Monte Carlo, Modèle d'Espace Latent.

Sommaire

Index des figures

1. Introduction

L'étude des réseaux sociaux est omniprésente, principalement lorsqu'il faut étudier les interactions dans un ensemble d'individus, dans différents contextes. Les exemples abondent dans toutes les sciences : dans l'industrie, avec la planification et l'optimisation des processus opérationnels ; en biologie, avec l'étude des systèmes complexes liés au génome humain ; dans l'économie, avec l'investigation des relations entre les nations, en termes de coopération économique, de droits de l'homme et d'actions militaires ; parmi tant d'autres.

Le domaine d'application qui illustre à la fois l'importance et les enjeux de l'étude des réseaux est celui des communications. La manière et la rapidité avec laquelle nous nous connectons sont actuellement un axe de développement important, car chaque individu établit des modes de relation différents selon son environnement. Par exemple, les différents rôles que remplissent les salariés d'une entreprise ne sont pas les mêmes que les relations interpersonnelles qui s'établissent en dehors de leur travail. Cependant, les deux manières d'interagir peuvent avoir un lien, même si elles sont différentes. Une portion de ces interactions se reflète dans les réseaux sociaux en ligne, tels que Twitter et Facebook, qui ont respectivement des approches et des objectifs différents, mais avec certaines structures en commun.

En étudiant les réseaux sociaux, il est généralement intéressant d'étudier la probabilité que deux acteurs interagissent entre eux, ainsi que leur « position sociale » dans le système. D'autant plus si les individus sont inscrits sur différents réseaux sociaux, et donc interagissent les uns avec les autres de différentes manières dans différents contextes. Ainsi, nous voulons étudier les dépendances structurelles à la fois "dans" et "entre" ces relations, étant donné qu'il existe différentes sources d'information dans lesquelles les mêmes acteurs du système sont impliqués.

Ce travail se structure comme suit: la Section 2 introduit les travaux déjà réalisés dans le domaine ; dans la Section 3, sont présentés plusieurs cas mettant en évidence des réseaux multiples ; dans la Section 4, les modèles spatiaux latents de distance sont étudiés ; dans la Section 5, un modèle pour la caractérisation de plusieurs réseaux est proposé, où les propriétés d'interchangeabilité sont également discutées avec l'élicitation des hyper-paramètres du modèle ; la Section 6 examine tous les aspects liés à la mise en œuvre du modèle ; la Section 7 présente les résultats correspondant aux exemples de motivation avec leur évaluation ; et enfin, dans la Section 8, les résultats sont discutés avec quelques alternatives pour les recherches futures.

1.1. Réseaux sociaux

Un réseau social (ou simplement réseau) désigne un ensemble de n acteurs (éléments ou individus) avec une variable $y_{i,i'}$ définie entre chacun d'eux, pour tout couple $i, i' = 1, \ldots, n$,

$i \neq i'$. Cette variable $y_{i,i'}$ caractérise l'interaction de deux individus i et i' dans un système particulier. Dans le cas le plus simple, la variable $y_{i,i'}$ est une variable dichotomique qui prend les valeurs 0 et 1 ; 0, s'il n'y a pas de lien entre les acteurs i et i' ; et 1, s'il y a effectivement une connexion.

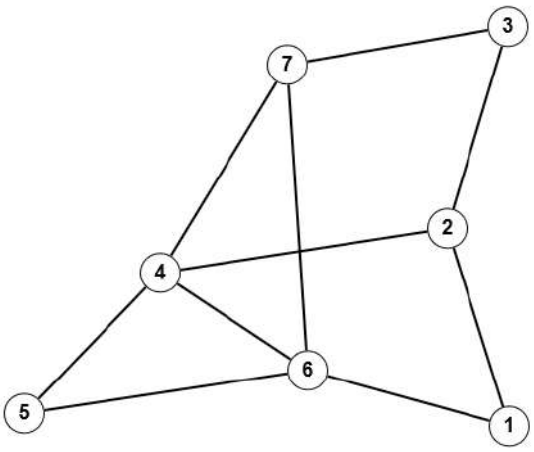

Figure 1: Exemple d'un graphe non dirigé avec $n=7$ acteurs.
Source: Élaboration personnelle

Les données peuvent être représentées par un graphe, qui est une collection de sommets (ou nœuds) et d'arêtes. Dans le contexte des réseaux sociaux, les acteurs correspondent aux sommets, tandis que les arêtes représentent les relations entre lesdits acteurs.
La Figure 1 illustre un réseau qui montre de $n=7$ sommets, qui peuvent représenter : les sénateurs d'un parti politique, membres d'une famille ou un ensemble de pays dans une région donnée.
Un réseau est dit non dirigé lorsque $y_{i,i'}=y_{i',i}$ pour toute paire de sommets, dans le cas contraire on dit que le réseau est dirigé. Un exemple de réseau dirigé est illustré par la Figure 2, où il y a une direction implicite dans les relations. On peut voir que le réseau de la Figure 1 est non-orienté, puisque les arêtes ne se réfèrent à aucune direction.
Par ailleurs, le lien entre deux sommets i et i' dans un réseau se note $i \sim i'$, ce qui implique qu'il y a une arête entre le sommet i et le sommet i'. Deux sommets connectés sont dits voisins, et le voisinage du sommet i, N_i, est l'ensemble de sommets voisins au sommet i.

$$N_i = \{ i' : i \sim i' \}$$

On peut présenter le graphe sous forme d'une matrice, où chaque interaction est représentée par un nombre (dans notre cas ce sera 1), indiquant qu'il y a interaction entre deux sommets, générant une matrice associée au réseau (ou matrice d'adjacence) Y, étant une matrice de $n \times n$ éléments $y_{i,i'}$ telle que :

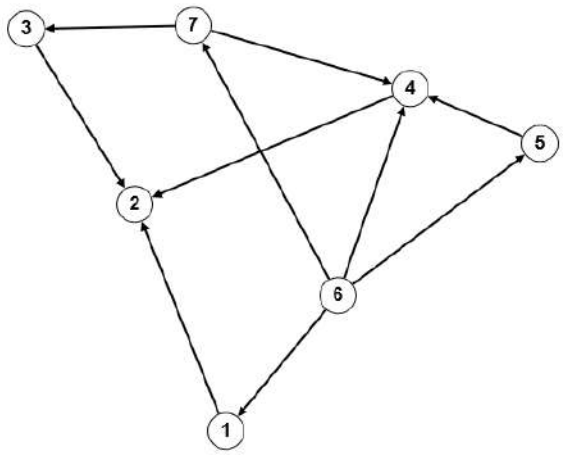

Figure 2: Exemple d'un réseau dirigé avec sept acteurs.
Source: Élaboration personnelle

$$y_{i,i'}=\begin{cases} 1 & si \;\; i\sim i',\;\; i\neq i', \\ 0 & autrement \end{cases}$$

On appelle sous-réseau de Y, un sous-ensemble Y' du réseau Y. Dans un réseau non orienté, sa matrice associée Y est symétrique, et peut être représentée par un vecteur y à $\binom{n}{2}$ composantes. Pour les illustrer, considérons le graphe de la Figure 1, dont la matrice d'adjacence est :

Y=

0	1	1	0	0	0	0
1	0	1	1	0	0	0
1	1	0	0	1	0	0
0	1	0	0	1	1	1
0	0	1	1	0	1	0
0	0	0	1	1	0	1
0	0	0	1	0	1	0

où $y=(1,1,0,0,0,0,1,1,0,0,0,0,1,0,0,1,1,1,1,0,1)'$ est la représentation vectorielle du réseau, ce ce qui permet de représenter les données de manière optimisée.
De la même manière, le degré d'un nœud i, d_i, est la le nombre de voisins (nœuds liés avec une arête). Cela s'écrit de la manière suivante :

$$d_i=|N_i|$$

La séquence de degré d'un graphe est un vecteur $(D_0,\ldots,D_M)$ où D_k est la le nombre de sommets dans le réseau de degré k, et M est la degré maximum observé dans la réseau. La distribution des degrés est donnée par $(D_0/n,\ldots,D_M/n)$, où n est la le nombre de nœuds du réseau.

Nous appellerons degré étendu du sommet i, e_i, la valeur suivante :

$$e_i = \sum_{i' : i \sim i'} |N_i / \{i\}|$$

En prenant en compte le réseau de la Figure 1, les degrés des sommets sont : $d_1=2$, $d_2=3$, $d_3=3$, $d_4=4$, $d_5=3$, $d_6=3$ et $d_7=2$. Ici, la séquence de degrés est $(0,2,4,1)$. Et par exemple, le degré étendu du nœud 3 est $e_3 = \sum_{i' \in N_3} d_{i'} - d_3 = 4$.

La longueur de la trajectoire v entre le sommet i et le sommet j est une suite d'indices $\{i_0, i_1, \ldots, i_{v+1}\}$ avec $i_0 = i$ et $i_{v+1} = j$, telle que :

$$\prod_{t=0}^{v} y_{i_t, i_{t+1}} = 1$$

On appelle la géodésique la trajectoire ayant la distance minimale entre le sommet i et i', $g_{i,i'}$. S'il n'y a pas de chemin, alors $g_{i,i'} = \infty$.

La séquence géodésique c'est $(G_0, \ldots, G_M)$, où G_K est la le nombre de nœuds pairs qui ont une distance géodésique minimale k, et M est la distance géodésique la plus grande. Aussi, on appelle triangle un cycle de trois longueurs dans réseau. Trois nœuds i, i', et k forment un triangle si $i \sim i'$, $i' \sim k$, et $k \sim i$.
La densité d'un réseau non dirigé est le rapport entre le nombre d'arêtes existantes et le nombre d'arêtes possibles. Par exemple, dans un réseau non orienté G, sans auto-boucles et sans arêtes multiples, la densité d'un sous-réseau $H = (V_H, E_H)$ est :

$$den(H) = \frac{|E_H|}{|V_H|(|V_H| - 1)/2}$$

Si la valeur $den(H)$ tend vers zéro, le réseau est dit creux, sinon si elle tend vers un, le réseau est dit dense.
La valeur $cl_T(G)$ est appelée la transitivité d'un réseau, elle est définie par :

$$cl_T(G) = \frac{3\tau_\Delta(G)}{\tau_3(G)}$$

où $\tau_\Delta(G)$ est le nombre de triangle dans le réseau G, et $\tau_3(G)$, le nombre de triplets connectés (c'est-à-dire un sous-réseau de trois sommets pour deux arêtes). La transitivité est une mesure de regroupement global, qui résume la fréquence relative avec laquelle les triplets proches se connectent pour former des triangles.
Un facteur important est la tendance à se connecter à des acteurs aux caractéristiques similaires, appelée assortativité, définie comme :

$$r_a = \frac{\sum_i f_{ii} - \sum_i f_{i+} f_{+i}}{1 - \sum_i f_{i+} f_{+i}}$$

où $f_{ii'}$ est la fraction des arêtes dans G qui unissent un sommet dans la i-ème catégorie avec un sommet dans la i-ième catégorie, tel que f_{i+} et f_{i+} désignent la i-ème ligne et colonne marginales, respectivement de la matrice résultante f.
La valeur r_a est comprise entre -1 et 1. Elle est égale à zéro lorsque le mélange dans le réseau n'est pas différent de celui obtenu par une affectation aléatoire des arêtes qui préservent la distribution des degrés marginaux.

2. Travaux existants

Un modèle statistique très populaire dans la littérature pour un seul réseau est le modèle de distance dans l'espace latent (Hoff et al., 2002). Selon ce modèle, les probabilités d'interaction dépendent de la distance entre les positions des acteurs dans un « espace social » latent, c'est-à-dire non observé. Cette formulation est très populaire car elle permet de modéliser directement des caractéristiques qui sont observées dans les réseaux sociaux, telles que la transitivité (l'ami d'un ami est un ami) et l'équilibre (l'ennemi d'un ami est un ennemi), par exemple. Des avancées significatives dans la modélisation des réseaux simples au moyen de variables latentes se trouvent dans les travaux de Nowicki et Snijders (2001), Hoff et al. (2002), Schweinberger et Snijders (2003), Hoff (2005), Handcock et al. (2007), Linkletter (2007), Krivitsky et Handcock (2008), Hoff (2008), Krivitsky et al. (2009), Hoff (2009), Li et al. (2011), Raftery et al. (2012), Minhas et al. (2018).

Dans ce travail, le modèle de distance de Hoff est étendu afin d'étudier conjointement le processus génératif de plusieurs réseaux transversaux (statiques). Dans cette optique, Gollini et Murphy (2016) et Salter-Townshend et McCormick (2017) apportent des contributions très importantes, permettant aux acteurs de jouer différents rôles dans différentes relations. Notre proposition s'appuie sur ces modèles, puisqu'elle maintient la vraisemblance multivariée de Bernoulli, mais structure davantage la distribution a priori correspondant au positionnement social de chaque acteur, ce qui permet de caractériser avec précision leurs caractéristiques au sein de chaque réseau et à travers les réseaux.

Du point de vue des espaces latents, outre les travaux pionniers de Salter-Townshend et McCormick (2017), d'autres alternatives pour étudier les réseaux sociaux multiples ont émergé au cours des dernières années. En connectomie cérébrale, Durant et al. (2018) proposent un modèle bayésien non paramétrique à travers un modèle de mélange qui réduit la dimensionnalité et intègre efficacement les informations des réseaux au sein de chaque composant du mélange. Puis Wang et al. (2017) proposent un modèle pour étudier les similitudes et les différences des matrices d'adjacence, basé sur une décomposition hiérarchique des valeurs singulières. À leur tour, D'Angelo et al., étendent les modèles d'espace latent dans d'autres contextes, en incorporant des effets spécifiques aux nœuds (D'Angelo et al., 2018) et des covariables spécifiques au lien D'Angelo et al. (2019), et enfin, en introduisant des groupes dans le cadre d'un mélange infini de distributions, D'Angelo et al. (2020). D'autres avancées importantes d'un point de vue fréquentiste sont disponibles dans Zhang (2020).

Il existe d'autres travaux en relation avec l'étude transversale des réseaux, comme ceux permettant la détection de communautés (par exemple, Han et al., 2014, Reyes et Rodriguez, 2016, Paul et al., 2016, Gao et al., 2019, Paez et al., 2019, et Paul et al., 2020), et l'évaluation de perception cognitive dans les structures sociales (par exemple, Swartz et al., 2015, Sosa, 2017, Sewell et al., 2020). Enfin, du point de vue dynamique, il existe une grande variété d'approches pour modéliser l'évolution d'un système dans le temps (e.g., Durante et Dunson, 2014, Hoff, 2014, Sewell et Chen, 2015, Sewell et Chen, 2016, Sewell et al., 2017, Gupta et al., 2018, Kim et al., 2018, Turnbull, 2020).

3. Quelques exemples de motivation

Les individus dans un environnement social peuvent être liés de différentes manières, générant ainsi un système de J réseaux sociaux sur même ensemble de I acteurs. Ci-dessous quelques exemples des réseaux binaires multiples non dirigés afin de montrer des scénarios où ce concept a sa place.

3.1. Salle de câblage

Roethlisberger et Dickson (2004)[1] ont recueilli des données auprès de $I=14$ employés de la salle de câblage de la compagnie Western Electric : deux inspecteurs, trois soudeurs et neuf opérateurs. Les auteurs ont recueilli $J=4$ types de relations, à savoir la participation à des jeux (Jeux), la participation à des dialogues à travers des fenêtres ouvertes (Dialogues), l'amitié (Amitié) et son comportement antagoniste (Hostilité). La Figure 3 et le Tableau 1 présentent une description des réseaux.

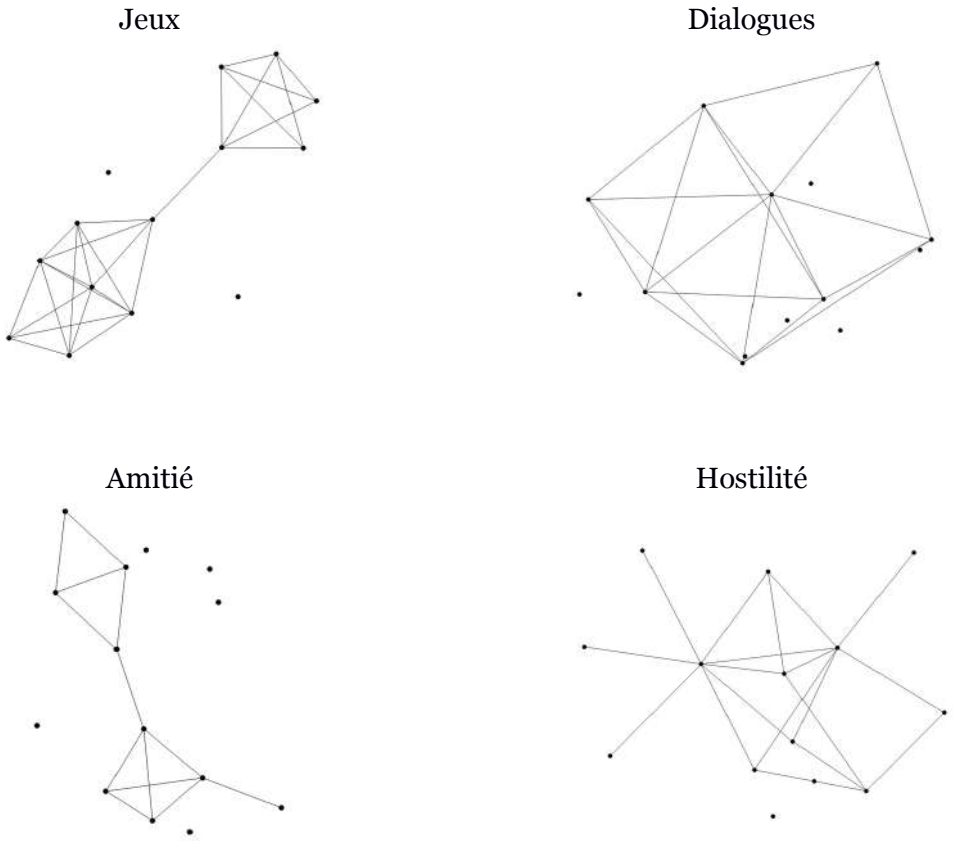

Figure 3: Graphes pour les jeux de données de la salle de câblage, affichés selon l'algorithme Kamada–Kawai.
Source: Capture d'écran du logiciel socnetv

1 La base de données est disponible sur: http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/wiring.dat

3.2. **Microfinance**

Banerjee et al. (2013)[2] ont rassemblé des données démographiques de plusieurs villes au Karnataka, au sud de l'Inde, afin d'étudier la micro-finance. Dans ce projet, un recensement des ménages et de leurs membres a été effectué, auquel un questionnaire détaillé a été établi sur les relations qu'ils avaient dans le village. Ces données ont été enregistrées pour $I=77$ maisons d'un village particulier, en termes de $J=3$ types de relations: échanges d'argent et d'équipement (échanges), de recommandations personnelles et médicales (conseils) et d'interactions sociales (sociabilité). La Figure 4 et le Tableau 1 montrent une description des réseaux.

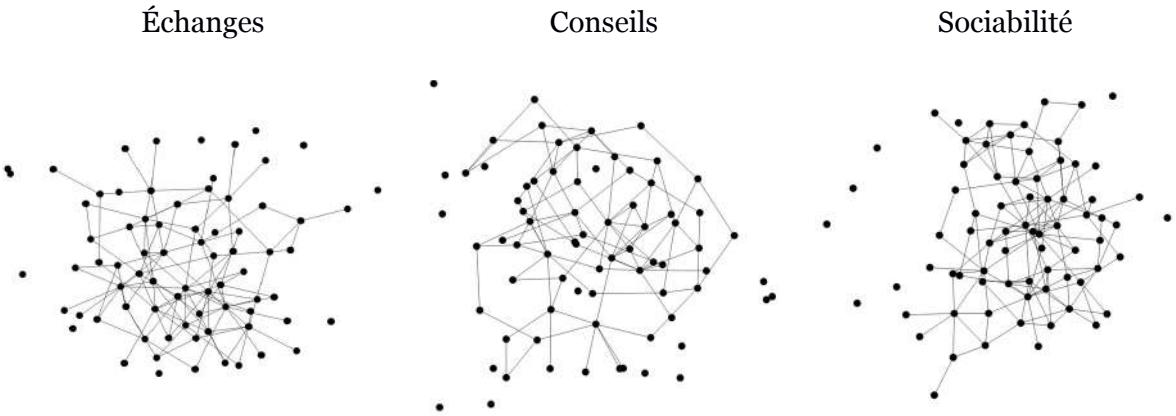

Figure 4: Graphes pour l'ensemble de données sur la microfinance.
Source: Élaboration personnelle

3.3. **Réseaux sociaux en ligne**

Huawei est une entreprise chinoise d'une grande importance au niveau mondial, et dans l'un de ses processus de recherche en technologie, elle a recueilli des informations sur les connexions sur un millier d'utilisateurs (anonymisés) qui utilisaient fréquemment $J=3$ réseaux sociaux, à savoir Facebook, Twitter et Instagram, dans afin d'orienter efficacement leurs processus marketing. L'ensemble de données est disponible sur https://www.kaggle.com/andrewlucci/huawei-social-network-data. Dans ce cas, une représentation a été réalisée avec un échantillon aléatoire de $I=100$ acteurs. La Figure 5 et le Tableau 1 présentent une description des réseaux.

2 La base de données est issue de https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/21538

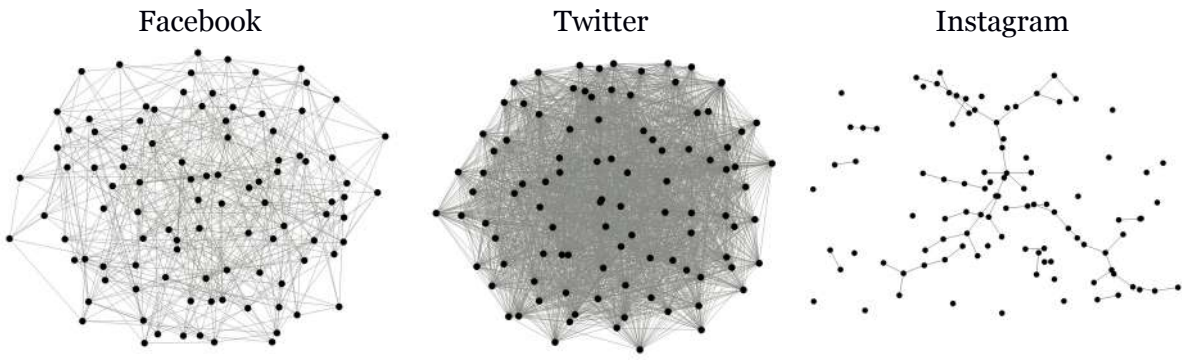

Figure 5: Graphes pour l'ensemble de données sur les réseaux sociaux en ligne.
Source: Élaboration personnelle

3.4. **Équipe de dirigeants**

Krackhardt(1987)[3] a recueilli des données auprès de la direction d'une entreprise de fabrication de machines de haute technologie pour évaluer les effets d'une intervention effectuée dans leur direction. $I=21$ employés ont été interrogés sur leurs amitiés perçues et leurs relations avec les autres, ce qui a donné $J=21$ réseaux, chacun associé à la perception de chaque employé. Un tel ensemble de données est appelé Cognitive Social Structure (CSS). Le Tableau 1 présente une description des réseaux.

Données	Abréviation	I	J	Densité	Transitivité	Assortativité
Salle de câblage	CÂBLAGE	14	4	0,217	0,591	-0,187
Ménages du sud de l'Inde	MAISONS	74	3	0,064	0,186	-0,215
Réseaux sociaux en ligne	SOCIAL	100	3	0,206	0,204	-0,560
Équipe de dirigeants	GESTION	21	21	0,124	0,337	-0,195

Tableau 1: Statistiques descriptives pour les exemples de motivation. Les valeurs de la densité, la transitivité et l'assortativité correspondent à la valeur moyenne de tous les réseaux.
Source: Élaboration personnelle

3 La base de données est issue de http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/krackfr.dat

4. Modèle de distance d'espace latent pour un réseau

Pour un réseau binaire, le modèle d'espace latent prend la forme d'une régression logistique. L'utilisation d'effets aléatoires dans le contexte de modèles linéaires généralisés est une alternative populaire pour modéliser des réseaux. Plus précisément, considérons un modèle dans lequel les $y_{i,i'}$ sont conditionnellement indépendants. La littérature offre les probabilités d'interaction suivantes :

$$Pr[y_{i,i'}=1|\beta,\gamma_{i,i'},x_{i,i'}]=expit(\alpha+x_{i,i'}^T\beta+\gamma_{i,i'}),i<i', \qquad (1)$$

Où la fonction $expit(x)=1/(1+e^{-x})$, étant l'inverse de la fonction *logit*, conserve l'ordre. α et β sont des effets fixes tandis que $\gamma_{i,i'}$ est un effet aléatoire spécifique non observé (i.e. latent). Nous bénéficions du vecteur prédicteur $x_{i,i'}=(x_{i,i',1},\ldots,x_{i,i',P})$, qui représente les covariables, s'il y en a, pour la paire de nœud associée à $y_{i,i'}$. α et β sont des paramètres de régression logistique. Plus spécifiquement, le vecteur $\beta=(\beta_1,\ldots,\beta_P)$ est le coefficient du vecteur $x_{i,i'}$. Il est nécessaire pour calculer le prédicteur linéaire $x_{i,i'}^T\beta=\sum_{p=1}^P\beta_p x_{i,i',p}$ associé à la paire (i,i'). α est une constante, l'ordonnée à l'origine. $\gamma_{i,i'}$ est associée à l'opposée d'une distance latente entre deux acteurs, comme nous allons l'expliquer. Autrement dit, la probabilité de l'existence d'une arête entre deux acteurs est déterminée par le vecteur $x_{i,i'}$ et la distance latente $\gamma_{i,i'}$.

Suivant les résultats de Hoover (1982) et Aldous (1985) (voir aussi Hoff, 2007, pour plus de détails), il est possible de montrer que si la matrice des effets aléatoires $\Gamma=[\gamma_{i,i'}]$ est conjointement interchangeable (voir Section 5.2), alors il existe une fonction symétrique $\alpha(\cdot,\cdot)$ telle que $\gamma_{i,i'}=\alpha(u_i,u_{i'})$, où $u_1,\ldots,u_I$ est un ensemble de vecteurs aléatoires indépendants latents. L'impact de telles variables dans (1) est largement dicté par la forme de la fonction $\alpha(\cdot,\cdot)$. C'est donc notamment par $\alpha(\cdot,\cdot)$ qu'il est possible de saisir d'autres caractéristiques pertinentes des données relationnelles. Plusieurs formules potentielles pour $\alpha(\cdot,\cdot)$ ont été explorées dans la littérature selon la catégorie d'espace latent.

Le modèle de distance de Hoff et al. (2002) suppose que chaque acteur i a une position $u_i=(u_{i,1},\ldots,u_{i,K})$ dans un espace social, ou espace des caractéristiques latentes, et que la probabilité d'un lien entre deux acteurs décroît lorsque les acteurs « s'éloignent » de l'espace social. À cette fin, les effets latents sont spécifiés par $\alpha(u_i,u_{i'})=-\|u_i-u_{i'}\|$, où $\|\cdot\|$ désigne la norme euclidienne. Les structures latentes basées sur la distance induisent la transitivité, qui est une caractéristique principale de nombreux réseaux sociaux réels. De plus, on observe que la modélisation des positions des acteurs dans un espace euclidien de faible dimension offre une alternative efficace pour représenter graphiquement des données relationnelles.

Étant donné un réseau binaire non orienté $Y=[y_{i,i'}]$, la vraisemblance (sans covariables) sous un modèle de distance est donnée par :

$$p(Y|\zeta,\{u_i\})=\prod_{i,i':i<i'}Ber(y_{i,i'}|expit(\zeta-\|u_i-u_{i'}\|)),$$

Où l'effet fixe $\zeta \in \mathbb{R}$ indique la propension moyenne à observer un lien entre deux acteurs donnés. Les variables $u_1, \ldots, u_I$ représentent les positions sociales non observées dans $\mathbb{R}^K$ des acteurs. On constate que le vraisemblance ne contient que $I(I-1)/2$ termes puisque Y correspond à la matrice d'adjacence d'un réseau non orienté.

Pour procéder à une analyse intégralement bayésienne et faire des inférences sur les paramètres du modèle, les distributions a priori pour ζ et chaque u_i doivent être spécifiées. Une option qui fonctionne généralement bien est $\zeta|\tau^2 \overset{iid}{\sim} N(0, \tau^2)$ et $u_i|\sigma^2 \overset{iid}{\sim} N(0, \sigma^2 \mathrm{I})$, où I (et non I) désigne la matrice identité ; on complète la formule de la distribution a priori avec $\tau^2 \sim IGam(a_\tau, b_\tau)$ et $\sigma^2 \sim IGam(a_\sigma, b_\sigma)$. Ainsi, l'ensemble des paramètres du modèle est $\breve{Y} = (\zeta, u_1, \ldots, u_I, \sigma^2, \tau^2)$, ce qui donne un total de $IK+3$ inconnues à estimer (quantités aléatoires), qui sont associées aux hyper-paramètres a_τ, b_τ, a_σ et b_σ. La Figure 6 montre la représentation du modèle de distance pour un seul réseau au moyen d'un *Directed Acyclic Graph* (DAG) et le Tableau 2 résume les relations entre hyper-paramètres, paramètres et la vraisemblance.

Une sélection précautionneuse des hyper-paramètres est primordiale pour garantir la performance appropriée du modèle. Pour cela, sont posés $a_\tau = 2$ et $b_\tau = 100$, qui place une distribution a priori diffuse pour ζ. De même, nous imitons une heuristique donnée dans Krivitsky et Handcock (2008, Sec. 2.4) en faisant a_σ et b_σ de sorte qu'a priori σ_2 est vaguement concentré (par exemple, $CV(\sigma^2) = 1$) autour de $E[\sigma^2] = I^{2/K}$, c'est-à-dire proportionnel au volume bidimensionnel d'une boule euclidienne de rayon $I^{1/K}$.

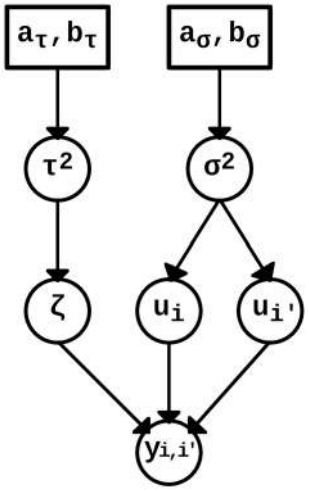

Figure 6: Représentation DAG du modèle de distance pour un seul réseau. Les cercles correspondent à des quantités aléatoires (paramètres), tandis que les rectangles à des quantités fixes (hyper-paramètres).
Source: Élaboration personnelle

Type de données	Données et relations		
Hyper-paramètres	a_τ, b_τ	a_σ, b_σ	
Paramètres	$\tau^2 \sim IGam(a_\tau, b_\tau)$	$\sigma^2 \sim IGam(a_\sigma, b_\sigma)$	
	$\zeta \overset{iid}{\sim} N(0, \tau^2)$	$u_i \overset{iid}{\sim} N(0, \sigma^2 I)$	$u_i \overset{iid}{\sim} N(0, \sigma^2 I)$
Vraisemblance	$p(Y \mid \zeta, \{u_i\}) = \prod_{i,i':i<i'} Ber(y_{i,i'} \mid expit(\zeta - \|u_i - u_{i'}\|))$		

Tableau 2: Récapitulatif des hyper-paramètres et paramètres nécessaires au modèle d'un réseau simple. Les paramètres sont spécifiés avec leurs distributions.
Source: Élaboration personnelle

5. Modèle de distance spatiale latente pour plusieurs réseaux

Cette section présente la proposition pour modéliser simultanément un ensemble de $J \geq 2$ réseaux sociaux binaires non orientés et non réflexifs $Y_1, \ldots, Y_J$, observés sur le même ensemble de I acteurs, où $Y_j = [y_{i,i',j}]$. À noter que dans ce cas $y_{i,i',j} \in \{0,1\}$ et $y_{i,i',j} = y_{i',i,j}$. Puisque chaque réseau contient des informations pertinentes sur le système social, au lieu de simplement ajuster des modèles de distance indépendants pour chaque réseau, le but de notre approche est d'extrapoler les informations entre les réseaux, afin d'obtenir de meilleurs résultats en termes de qualité d'ajustement et de prédiction au sein de chaque réseau.

5.1. La modélisation

Notre proposition est basée sur une extension hiérarchique en plusieurs étapes du modèle pour un seul réseau discuté dans la Section 4. Pour cela, on suppose que les connexions sont conditionnellement indépendantes avec la distribution de Bernoulli, ce qui se traduit par une vraisemblance donnée par :

$$p(Y|\{\zeta_j\},\{u_{i,j}\}) = \prod_j \prod_{i,i':i<i'} Ber(y_{i,i',j}|expit(\zeta_j - \|u_{i,j} - u_{i',j}\|))$$

Où, par extension de la vraisemblance donnée dans la Section 4, $\zeta_j \in \mathbb{R}$ est un effet fixe qui représente la propension moyenne à observer une connexion dans le réseau j entre deux acteurs donnés. Le vecteur $u_{i,j} = (u_{i,j,1}, \ldots, u_{i,j,K})$ désigne la position sociale de l'acteur *i* dans le réseau j, défini dans un espace euclidien de dimension K. Dans ce contexte, l'interprétation des positions est identique : si les caractéristiques latentes des individus *i* et *i'* du réseau j sont « plus éloignées » dans l'espace social, alors la distance $\|u_{i,j} - u_{i',j}\|$ augmente, et donc la probabilité d'observer un lien entre les acteurs *i* et *i'* diminue.
Suivant la formule standard des modèles de distance, on assigne aux positions sociales $u_{i,j}$ des distributions normales conditionnellement indépendantes, mais cette fois avec une moyenne et une variance sujettes à l'état de chaque individu $u_{i,j}|\eta_i, \sigma_j^2 \overset{ind}{\sim} N(\eta_i, \sigma_j^2 \mathrm{I})$. La moyenne $\eta_i = (\eta_{i,1}, \ldots, \eta_{i,K})$ peut être interprétée comme la position sociale « globale » de l'acteur *i* par rapport à l'ensemble des relations dans tous les réseaux du système. À présent, en établissant que les moyennes globales η_i et les composantes de la variance σ_j^2, bénéficient d'une distribution commune, $\eta_i|\theta, \kappa^2 \overset{iid}{\sim} N(\theta, \kappa^2 \mathrm{I})$ et $\sigma_j^2|\alpha_\sigma, \beta_\sigma \overset{iid}{\sim} IGam(\alpha_\sigma, \beta_\sigma)$ respectivement, alors il est possible de capturer des similitudes entre les acteurs, et ainsi, partager les informations à travers les réseaux. Pour compléter la formule hiérarchique de ce niveau, on a $\theta \sim N(\mu_0, \Sigma_0)$, $\kappa^2 \sim IGam(a_\kappa, b_\kappa)$, $\alpha_\sigma \sim Unif(c_\sigma, d_\sigma)$, et $\beta_\sigma \sim Gam(a_\sigma, b_\sigma)$. Enfin, en suivant une idée similaire, on attribue aux effets fixes des probabilités d'interaction ζ_j une distribution a priori hiérarchique conjuguée, $\zeta_j|\mu_\zeta, \tau_\zeta^2 \sim N(\mu_\zeta, \tau_\zeta^2)$, avec $\mu_\zeta \sim N(\mu_0, \sigma_0^2)$ et $\tau_\zeta^2 \sim IGam(a_\zeta, b_\zeta)$.
Ainsi, l'ensemble des paramètres du modèle est le suivant :

$$\breve{Y} \equiv \breve{Y}_{I,J,K} = (\zeta_1, \ldots, \zeta_J, u_{1,1}, \ldots, u_{I,J}, \eta_1, \ldots \eta_I, \sigma_{1,\ldots}^2, \sigma_J^2, \mu_\zeta, \tau_\zeta^2, \theta, \kappa^{2,} \alpha_\sigma, \beta_\sigma)$$

Il contient un total de $IK(J+1)+2J+K+5$ inconnues à estimer, associées aux hyper-paramètres μ_0, Σ_0, a_κ, b_κ, μ_0, σ_0^2, a_ζ, b_ζ, a_σ, b_σ, c_σ et d_σ. La Figure 7 montre la représentation du modèle de distance pour des réseaux multiples au moyen d'un DAG. Dans cette représentation, les trois niveaux du modèle sont évidentes, tout comme dans le résumé fournit par le Tableau 3.

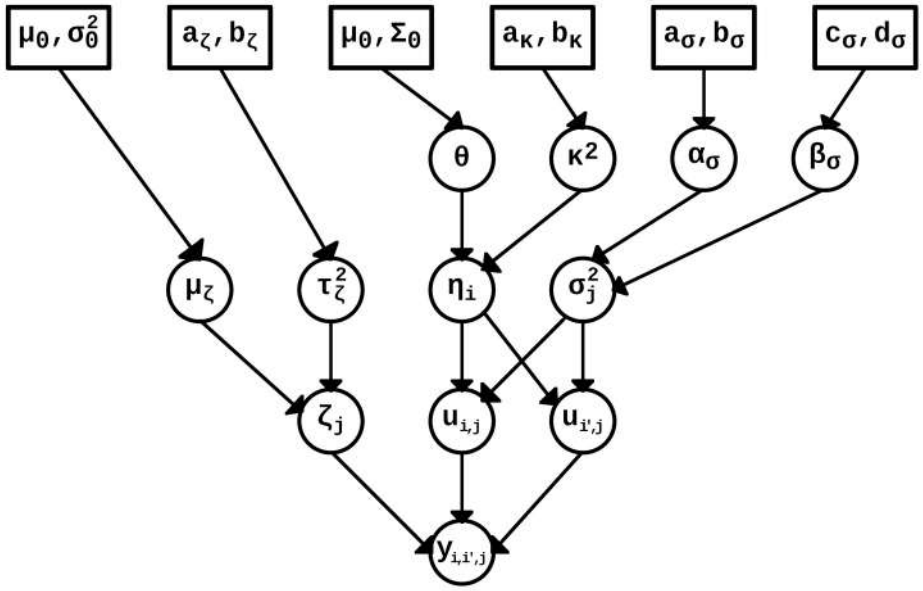

Figure 7: Représentation DAG du modèle de distance pour réseaux multiples. Les cercles correspondent à des valeurs aléatoires (paramètres), alors que les rectangles correspondent à des valeurs fixes (hyper-paramètres).
Source: Élaboration personnelle

Hyper-paramètres	Paramètres et leurs distributions			Vraisemblance
μ_0, σ_0^2	$\mu_\zeta \sim N(\mu_0, \sigma_0^2)$		$\zeta_j \sim N(\mu_\zeta, \tau_\zeta^2)$	$p(Y\|\{\zeta_j\},\{u_{i,j}\})$
a_ζ, b_ζ	$\tau_\zeta^2 \sim IGam(a_\zeta, b_\zeta)$			
μ_0, Σ_0	$\theta \sim N(\mu_0, \Sigma_0)$	$\eta_i \overset{iid}{\sim} N(\theta, \kappa^2 I)$	$u_{i,j} \overset{ind}{\sim} N(\eta_i, \sigma_j^2 I)$	
a_κ, b_κ	$\kappa^2 \sim IGam(a_\kappa, b_\kappa)$			
c_σ, d_σ	$\alpha_\sigma \sim Unif(c_\sigma, d_\sigma)$	$\sigma_j^2 \overset{iid}{\sim} IGam(\alpha_\sigma, \beta_\sigma)$	$u_{i',j} \overset{ind}{\sim} N(\eta_i', \sigma_j^2 I)$	
a_σ, b_σ	$\beta_\sigma \sim Gam(a_\sigma, b_\sigma)$			

Tableau 3: Récapitulatif des hyper-paramètres et paramètres nécessaires au modèle de réseaux multiples. La formule de la vraisemblance n'est pas affichée ici.
Source: Élaboration personnelle

5.2. **Interchangeabilité**

Aldous (1985) étend les notions d'interchangeabilité de De Finetti aux matrices bidimensionnelles, grâce à l'introduction des concepts d'interchangeabilité séparée (lorsque la distribution conjointe d'une matrice bidimensionnelle de variables aléatoires est invariante et indépendante des permutations arbitraires de lignes et de colonnes) ; et d'interchangeabilité conjointe (lorsque la distribution conjointe d'une matrice bidimensionnelle de variables aléatoires n'est invariante que si la même permutation arbitraire est appliquée aux lignes et aux colonnes). L'interchangeabilité conjointe implique l'interchangeabilité séparée, mais pas l'inverse. Ces notions peuvent facilement être étendues à des matrices de dimension supérieure, en considérant par exemple toutes les paires de dimensions possibles.
Dans le contexte des réseaux multiples, π_1 et π_2 sont deux permutations de $\{1,\ldots,I\}$, et π_3 une permutation de $\{1,\ldots,J\}$. Sous la version du modèle présenté dans la Section 5.1, la distribution marginale conjointe des observations $\{y_{i,i',j}\}$ est la même que la distribution de $\{y_{\pi_1(i),\pi_2(i'),\pi_3(j)}\}$ seulement si $\pi_1=\pi_2$. En d'autres termes, le modèle définit une distribution de probabilité interchangeable jointe pour les données issues de réseaux multiples. L'interchangeabilité jointe (à la place d'une forme plus faible d'interchangeabilité) est particulièrement intéressante dans ce contexte, car tous les indices *i* et *i'* (et potentiellement *j*) renvoient au même ensemble d'acteurs.

5.3. **Élicitation des hyper-paramètres**

Au préalable, une sélection rigoureuse des hyper-paramètres est fondamentale pour garantir les performances adéquates du modèle. Ainsi, pour la distribution a priori de θ, on pose $\mu_0=0$ et $\Sigma_0=100\,\mathrm{I}$, qui se concentre a priori sur l'origine aux moyennes globales η_i de manière non informative. Dans l'approche du modèle latent pour un seul réseau (Section 4) il est indiqué que l'implémentation du modèle se fait avec des hyper-paramètres qui génèrent une distribution a priori non informative, afin d'éviter la concentration dégénérée des positions latentes en un point unique de l'espace social. Pour cette raison, il est recommandé de suivre l'heuristique donnée dans Krivitsky et Handcock (2008, Sec. 2.4), implémentable avec la bibliothèque R *latentnet*, afin de ne pas se concentrer a priori sur la distribution des positions latentes des sommets, avant d'observer les réseaux. Ce même concept a été mis en œuvre dans le cas de réseaux multiples.
Suivant la même idée pour les effets fixes ζ_j, on pose $\mu_0=0$ et $\sigma_0^2=100$. Ensuite, pour la composante de variance κ^2, on établit une loi de distribution a priori de moyenne finie $I^{2/K}$, mais de variance infinie, avec $a_\kappa=2$ et $b_\kappa=I^{2/K}$; ceci permet de faire varier les η_i de manière diffuse selon l'heuristique donnée dans la Section 4. De même, pour la composante de variance τ^2, on pose $a_\zeta=2$ et $b_\zeta=1$, ce qui permet a priori de faire varier raisonnablement les ζ_j dans une plage de valeurs adéquate. Enfin, on pose $a_\sigma=1$, $b_\sigma=J^{-1/K}$, $c_\sigma=0$ et $d_\sigma=10$ afin de laisser les caractéristiques latentes de chacun des réseaux varier de manière diffuse dans un espace raisonnable, en fonction de la quantité de réseaux.

6. Calcul

Pour une dimension latente K fixée, on peut explorer la distribution postérieure des paramètres en utilisant des méthodes de Markov Chain Monte Carlo (MCMC ; e.g., Gamerman and Lopes, 2006), où la distribution postérieure est approximée en utilisant des échantillons dépendants, mais à peu près identique, $\breve{Y}^{(1)},\ldots,\breve{Y}^{(B)}$ de la distribution postérieure $p(\breve{Y}|Y) \propto p(Y|\breve{Y})p(\breve{Y})$, où $\breve{Y}$ est l'ensemble des paramètres du modèle donné dans la Section 5.1. Les estimations ponctuelles et les intervalles de crédibilité peuvent être approximés à partir des distributions empiriques fournies par les échantillons.
Dans notre algorithme, lorsque cela est possible, nous prélevons des échantillons à partir des distributions conditionnelles complètes comme dans un échantillonneur de Gibbs habituel ; sinon, on utilise des versions adaptatives avec des étapes de Metropolis-Hastings (par exemple, Haario et al., 2001). Des détails sur l'algorithme mis en œuvre ici peuvent être trouvés dans l'annexe A.

6.1. Identifiabilité

Les modèles de distance sont invariables aux rotations et aux reflets de l'espace social. En effet, pour toute matrice orthogonale Q de taille $K \times K$, la vraisemblance associée avec le reparamétrage $\tilde{u}_{i,j}=Qu_{i,j}$ est indépendante de Q, étant donné que $\|\tilde{u}_{i,j}-\tilde{u}_{i,j'}\| = \|u_{i,j}-u_{i,j'}\|$. Dans le modèle présent, les paramètres qui ne sont pas identifiables sont les positions latentes $u_{i,j}$, parce que les modèles de distance entre les ensembles de points d'un espace euclidien sont invariants par mouvements rigides, car associés à chaque tenseur de position latente $U=[u_{i,j}]$. Il existe un nombre infini de positions qui produisent le même vraisemblance que celle du modèle proposé. Voir la Section 6.2 pour plus détails.
Cependant, comme le prédicteur linéaire $\eta_{i,i',j} = \zeta_j - \|u_{i,j}-u_{i',j}\|$ est parfaitement identifiable, et par conséquent, les probabilités d'interaction $expit(\eta_{i,i',j})$ restent identifiables grâce à l'injectivité de la fonction $expit$, on conclu que le modèle proposé est complètement identifiable. Voir (Hoff et al., 2002, p. 1091) pour plus de détails.
Ce problème est résolu en utilisant une approche de transformation de paramètres similaire à celle décrite dans Hoff et al. (2002, article 3). En particulier, l'inférence est restreinte à une classe particulière de positions latentes, au moyen d'une transformation des échantillons postérieurs $u_{i,j}^{(b)}$ en un système de coordonnées partagé. Ainsi, pour chaque échantillon $\breve{Y}^{(b)}$, une matrice de transformation orthogonale $Q^{(b)}$ est obtenue, et elle minimise la distance de Procuste,

$$\tilde{Q}^{(b)}=argmin_{Q\in S^k} tr\{(W^{(1)}-W^{(b)}Q)^T(W^{(1)}-W^{(b)}Q)\},$$

où S^K désigne l'ensemble des matrices orthogonales de taille $K \times K$ et $W^{(b)}$ est une matrice rectangulaire de taille $I \times K$ dont les lignes correspondent aux moyennes globales $\eta_1^{(b)},\ldots,\eta_I^{(b)}$. Ce problème d'optimisation peut être facilement résolu à l'aide d'une

décomposition en valeurs singulières (par exemple, Borg et Grönen, 2005, Sec. 20.2). Une fois les matrices obtenues $\widetilde{Q}^{(1)},\ldots,\widetilde{Q}^{(B)}$, l'inférence postérieure pour les positions latentes est basée directement sur $\widetilde{u}_{i,j}^{(b)}=\widetilde{Q}^{(b)}u_{i,j}$ et $\widetilde{\eta}_i^{(b)}=\widetilde{Q}^{(b)}\eta_i^{(b)}$.

6.2. **Sélection de la dimension latente** K

La valeur $K=2$ est populaire dans la littérature sur la modélisation de l'espace latent, car elle simplifie la visualisation et facilite ainsi la description des relations sociales. Cependant, notre objectif va au-delà d'une simple description des réseaux, de sorte que la valeur de K joue un rôle critique dans les résultats. Certaines méthodologies pour sélectionner correctement la dimension de l'espace social sont discutées ci-dessous.
La littérature sur les réseaux s'est largement concentrée sur le critère d'information bayésien (BIC ; par exemple, Hoff, 2005, Handcock et al., 2007). Cependant, le BIC est souvent inapproprié pour les modèles multiniveaux car la structure hiérarchique implique que le nombre effectif de paramètres soit généralement inférieur au nombre réel de paramètres dans la vraisemblance. Une alternative au BIC est le Watanabe-Akaike Information Criterion (WAIC ; Watanabe, 2010, Gelman et al., 2013),

$$WAIC(K)=-2\sum_j \sum_{i,i':i<i'} \log E[\,p(y_{i,i',j}|\breve{Y})]+2\,p_{WAIC}$$

Où

$$p_{WAIC}=2\sum_j \sum_{i,i':i<i'} \left(logE[\,p(y_{i,i',j}|\breve{Y})]-E[\log p(y_{i,i',j}|\breve{Y})]\right)$$

correspond à la complexité du modèle (le nombre effectif de paramètres), et $\breve{Y}$ est l'ensemble de paramètres en supposant que la dimension de l'espace social soit K. Les valeurs attendues dans ces expressions sont calculées par rapport à la distribution postérieure $p(\breve{Y}|Y)$, qui peut être approximée à l'aide des échantillons $\breve{Y}^{(1)},\ldots,\breve{Y}^{(B)}$ de l'algorithme basé sur MCMC.

7. Illustrations

7.1. Salle de câblage

On analyse dans cette partie les réseaux multiples de la salle de câblage introduits dans la Section 3, en utilisant notre proposition. Les résultats présentés ci-dessous sont basés sur $B=20\,000$ échantillons de la distribution postérieure obtenue après réduction des chaînes de Markov originales toutes les 25 observations et une période de chauffe (warm-up period) de 20 000 itérations. La convergence a été surveillée en suivant la variabilité de la distribution conjointe des données et des paramètres du modèle à l'aide de la procédure à plusieurs volets donnée dans Gelman et al. (1992). Par exemple, le panneau de droite de la Figure 8 montre la chaîne de log-vraisemblance associée à la valeur de la dimension latente qui optimise le critère d'information WAIC, où l'on observe qu'il n'y a aucune preuve apparente d'absence de convergence.

Dimension de l'espace social

La Figure 8 montre les valeurs WAIC associées à des modèles ajustés avec différentes dimensions K de l'espace social. Ce critère privilégie un choix de $K=4$ (WAIC = 201,46), qui est la valeur utilisée dans toutes les analyses de cet ensemble de données. Ainsi, notre modèle utilise 297 paramètres avec $K=4$.

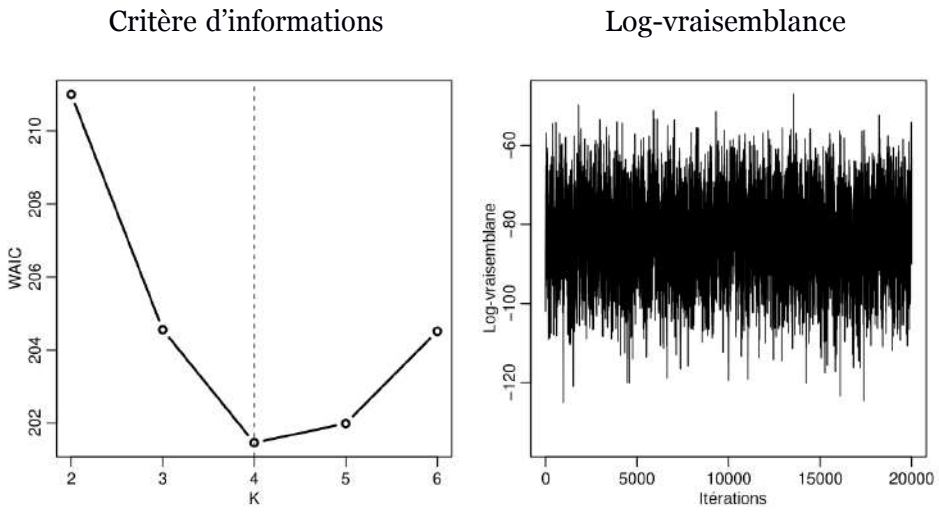

Figure 8: À gauche, les valeurs WAIC pour sélectionner la dimension K de l'espace social pour l'analyse des données de la salle de câblage, en ajustant le modèle proposé dans la Section 5. À droite, la chaîne de Markov de la log-vraisemblance associée à la valeur de K qui optimise le critère d'information WAIC.
Source: Élaboration personnelle

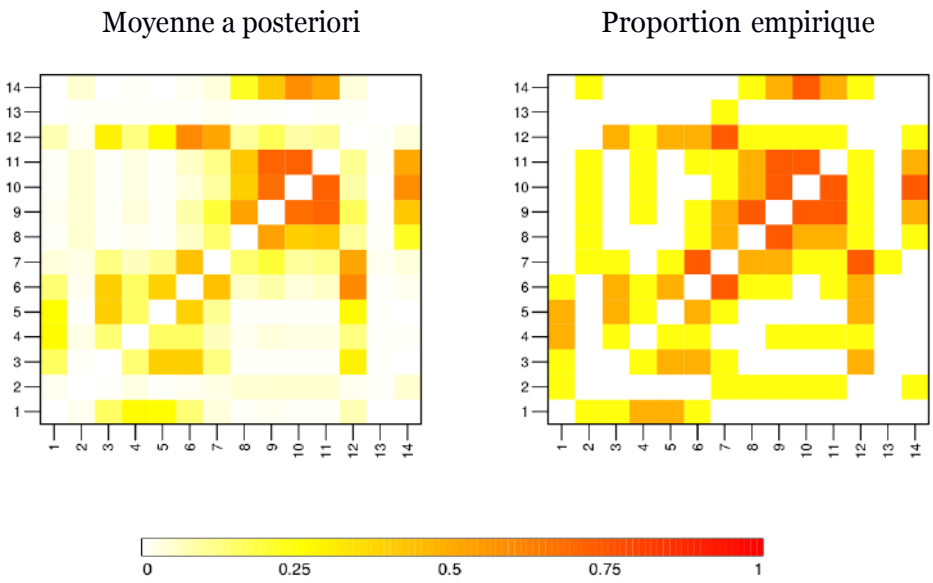

Figure 9: Deux estimations de la probabilité consensus de l'interaction entre deux acteurs. La panneau la gauche fournit la moyenne postérieure correspondante sous notre modèle, tandis que le panneau de droite montre la proportion de liens observés à travers les réseaux.
Source: Élaboration personnelle

Réseau consensuel

Notre modèle peut être utilisé pour produire une estimation d'un réseau de "consensus", qui contient l'information de tous les réseaux dans un seul réseau. En particulier, la probabilité consensuelle d'observer un lien entre les acteurs i et i' peut être estimée avec $\vartheta_{i,i'} = expit(\mu_\zeta - \|\eta_i - \eta_{i'}\|)$. La Figure 9 présente la matrice des moyennes a posteriori,

$$E[\vartheta_{i,i'}|Y] = \frac{1}{B}\sum_{b=1}^{B} expit(\mu_\zeta^{(b)} - \|\eta_i^{(b)} - \eta_{i'}^{(b)}\|)$$

avec une carte thermique de la proportion de liens observés à travers les réseaux, $\frac{1}{J}\sum_{j=1}^{I} y_{i,i',j}$. Même si l'estimation que fournit notre modèle est moins « dense » que la proportion empirique, on observe que les estimations sont très similaires. Ceci suggère aussi que le modèle caractérise correctement le processus de génération des données.

Projections dans l'espace social

Les positions latentes $u_{i,j}$ dans l'espace social sont un outil puissant pour décrire les interactions sociales. La première ligne de la Figure 10 montre les coordonnées selon les deux dimensions de plus grande variabilité des positions $u_{i,j}$ pour chacun des quatre réseaux. On observe que les dynamiques sociales de Jeux et d'Amitié sont très similaires, sauf que dans Jeux les acteurs 1 et 4, ils ont la même position sociale ; de même, dans ces deux relations, deux groupes d'individus sont nettement différenciés. La relation Dialogues a également une certaine ressemblance avec Jeux et Amitié, même si la différenciation des groupes est moins marquée. De plus, les caractéristiques de l'acteur 3 ne correspondent pas à celles des autres acteurs. D'un autre côté, les acteurs qui ont des positions proches dans

Amitié, ont généralement des positions éloignées dans Hostilité, et vice-versa. Cet effet est particulièrement clair entre les acteurs $\{1;4\}$ et $\{10;14\}$. La description des rôles sociaux peut également être effectuée au moyen des positions moyennes η_i.
Comme précédemment constaté, les positions latentes $u_{i,j}$ permettent aussi de distinguer des groupes d'acteurs qui remplissent des rôles sociaux similaires. Afin d'identifier ces groupes, il est possible d'appliquer une technique de clustering non supervisée (e.g., clustering hiérarchique, k-means), ou au contraire, d'inclure directement dans le modèle un ensemble de paramètres qui affectent les acteurs aux groupes. Cette dernière alternative est préférable car elle permet de quantifier directement l'incertitude liée au regroupement ainsi que sa relation avec les autres paramètres du modèle (voir la Section 8 pour plus de détails).

Ajustement du modèle

Ici, on évalue le modèle en utilisant des métriques dans l'échantillon. Premièrement, on compare les matrices d'adjacence observées, $Y_j=[y_{i,i',j}]$ avec les moyennes postérieures correspondantes aux probabilités d'interaction $\hat{\Theta}_j=[\hat{\theta}_{i,i',j}]$, où

$$\hat{\theta}_{i,i',j}=E[expit(\zeta_j-\|u_{i,j}-u_{i',j}\|)|Y]=\frac{1}{B}\sum_{b=1}^{B} expit(\zeta_j^{(b)}-\|u_{i,j}^{(b)}-u_{i',j}^{(b)}\|)$$

La Figure 10 montre les probabilités d'interaction estimées. En général, on observe que les probabilités estimées coïncident avec les relations observées dans tous les réseaux, ce qui constitue une preuve que le modèle s'adapte correctement aux données.
À présent, suivant Gelman et al. (2013, Ch. 6), on explore plus en détail la qualité de l'ajustement du modèle en générant des répliques de l'ensemble de données à partir des paramètres de la chaîne de Markov du modèle ajusté, puis en calculant un ensemble de statistiques récapitulatives pour chaque échantillon, dont les distributions sont ensuite comparées aux valeurs observées dans l'échantillon d'origine. La Figure 11 montre des résumés de la distribution empirique des statistiques de test sélectionnées (assortativité, transitivité et densité) pour chaque réseau (voir Kolaczyk et Csárdi, 2014 pour plus de détails sur les statistiques de test). Le modèle saisit adéquatement les caractéristiques structurelles des données étant donné que toutes les valeurs observées appartiennent aux intervalles de crédibilité. Les résultats avec le nombre de triangles, la distance géodésique, la eigen-centralité moyenne, le degré moyen et l'écart type du degré (non représentés ici) sont équivalents.

7.2. Évaluation de la prédiction et de la qualité d'ajustement

Afin d'évaluer la capacité du modèle proposé (que l'on appellera ENSEMBLE) à prédire les liens perdus, cette section évalue son pouvoir prédictif hors échantillon par le biais d'une expérience de validation croisée, en utilisant les quatre systèmes multi-réseaux présentés dans la Section 3 (voir le Tableau 1 pour plus de détails). De plus, comme point de

référence, le modèle de distance pour un seul réseau discuté dans la Section 4, ajusté indépendamment sur chacun des réseaux (que l'on appellera INDÉPENDANT), est également considéré pour effectuer une comparaison.

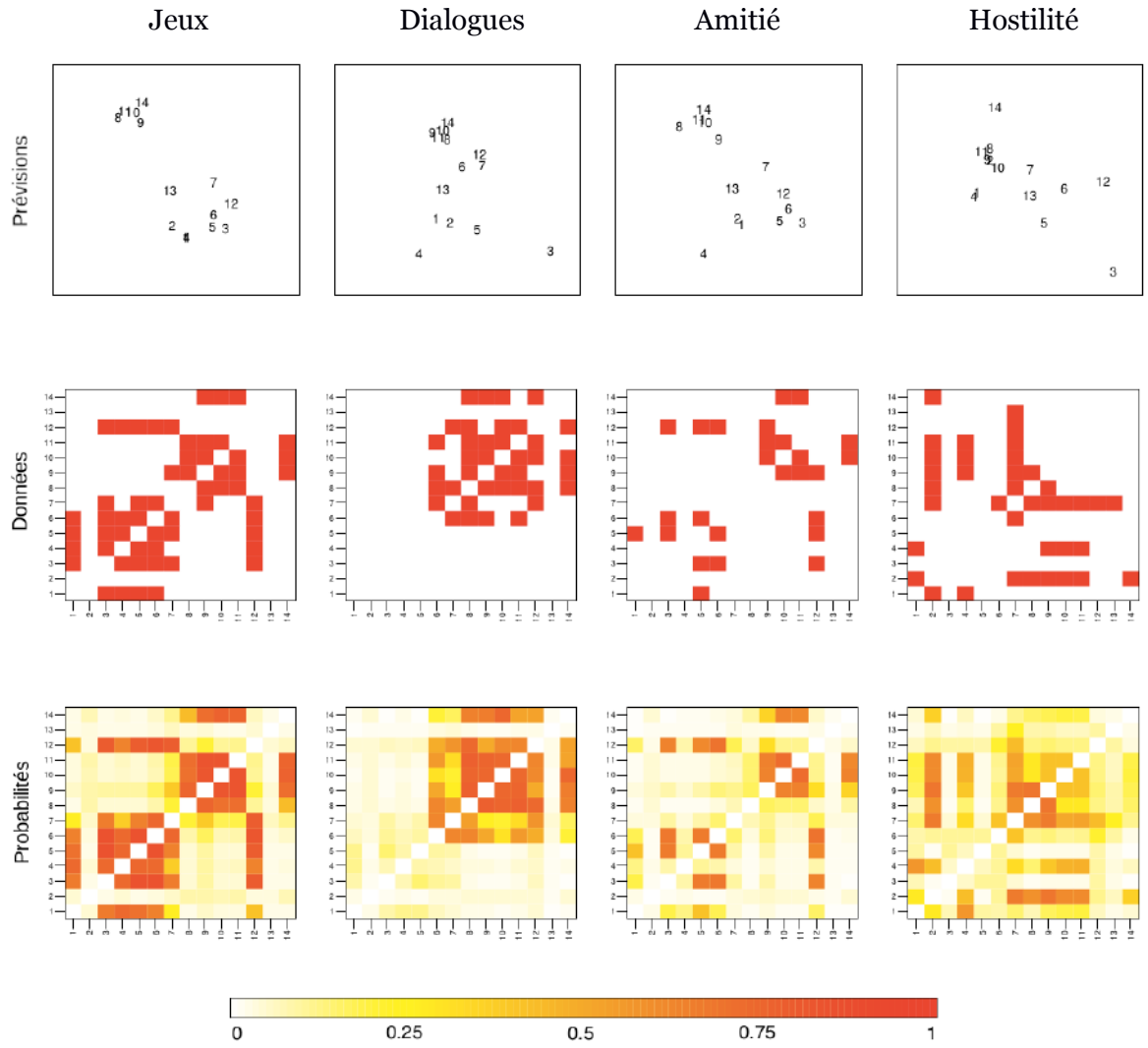

Figure 10: En haut, les positions des variables latentes $u_{i,j}$ en deux dimensions avec une meilleure variabilité ; au milieu, les matrices d'adjacence Y_j ; et en bas, les matrices de probabilité d'interaction estimées, $\hat{\Theta}_i$. Le tout pour l'ensemble de données de la salle de câblage.
Source: Élaboration personnelle

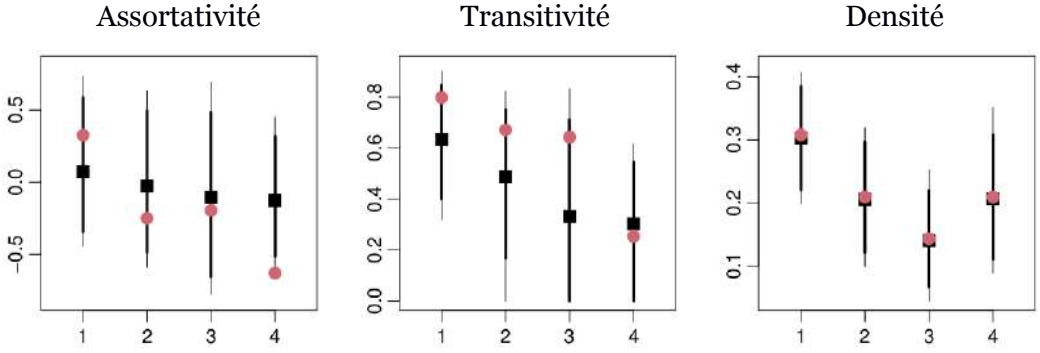

Figure 11: Intervalles de crédibilité à 95 % (trait épais) et 99 % (trait fin), moyenne postérieure (carré noir), et valeur observée (cercle rouge) de la distribution empirique de la densité, la transitivité, et de l'assortativité de 20 000 répliques de l'ensemble des données.
Source: Élaboration personnelle

Pour autant, pour chaque combinaison de modèle (ENSEMBLE, INDÉPENDANT), d'ensemble des données (CÂBLAGE, MAISONS, SOCIAL, GESTION), et la dimension latente (valeur de $K \in [\![1;6]\!]$ qui minimise le WAIC), se réalise une expérience de validation croisée de la manière suivante. D'abord, on divise les données aléatoirement dans cinq ensembles de même taille environ. Puis, pour chaque ensemble S, on ajuste le modèle conditionnel en $D=\{y_{i,i',j}:(i,i',j)\notin S\}$, et pour chaque y^* dans S, on calcule $E[y^*|D]$, la moyenne prédictive postérieure de y^* en utilisant toutes les données n'étant pas dans S. Ensuite, en utilisant de telles prédictions, on construit un classificateur binaire pour obtenir la courbe Receiver Operating Characteristic (ROC) correspondante. Enfin, on quantifie la capacité prédictive de chaque courbe ROC grâce à l'Area Under the Curve (AUC). Dans ce contexte, l'AUC est une mesure de la capacité d'un modèle à prédire les liens perdus. Plus l'AUC est élevée, meilleur est le modèle pour prédire les 0 comme des 0 et les 1 comme des 1. Dans tous les cas, les inférences sont basées sur des échantillons de 20 000 échantillons de la distribution postérieure obtenues après les réductions des chaînes de Markov originales chaque 25 observations et une période de chauffe de 20 000 itérations.

Mesure	AUC		WAIC		BIC	
Données	ENS.	INDE.	ENS.	INDE.	ENS.	INDE.
CÂBLAGE	**0,8827**	0,8886	**199,20**	216,48	**2 433,00**	1 062,82
MAISONS	**0,8866**	0,7765	**2 513,41**	3 442,38	**19 032,71**	22 191,42
SOCIAL	**0,8322**	0,8340	**10 577,20**	10 855,31	**29 934,11**	17 187,44
GESTION	**0,9036**	0,7974	**1 505,73**	2 267,71	**25 299,21**	15 116,49

Tableau 4: Valeurs de l'AUC, du WAIC et du BIC pour évaluer, respectivement le pouvoir prédictif et la qualité d'ajustement d'ENSEMBLE et INDÉPENDANT, en utilisant les réseaux multiples du Tableau 1. L'AUC correspond à la valeur moyenne des cinq expériences de validation croisée. Alors que pour le WAIC et le BIC, il correspond à la plus petite valeur de K comprise entre 1 et 6 qui minimise le WAIC.
Source: Élaboration personnelle

Les résultats de la validation croisée sont présentés dans le Tableau 4. Ces résultats suggèrent qu'en termes prédictifs, ENSEMBLE est égal ou plus compétitif qu'INDEPENDENT, ce qui est particulièrement évident dans MAISON et GESTION où la capacité à prédire les liens perdus d'ENSEMBLE est nettement supérieure.
Finalement, pour évaluer la qualité de l'ajustement de chaque modèle, le WAIC est calculé (voir la Section 5 pour plus de détails) pour tous les ensembles de données. Le Tableau 4 présente les résultats pour la plus petite valeur pour $K \in [\![1;6]\!]$ qui minimise le critère d'information. Dans ce cas, ENSEMBLE présente une meilleure qualité d'ajustement dans tous les cas, et constitue donc une meilleure option pour caractériser le processus aléatoire qui donne lieu aux ensembles de données.

8. Discussion

Ce travail présente une nouvelle approche pour modéliser un système de réseaux multiples sur le même ensemble d'acteurs avec une méthode qui encourage le flux d'informations entre les réseaux, contrairement à une caractérisation indépendante de chacun d'eux. Notre proposition est basée sur une extension hiérarchique naturelle d'un modèle de distance d'espace latent et permet d'étudier simultanément les positions sociales des acteurs à l'intérieur et entre les réseaux. De plus, nos expériences fournissent des preuves empiriques suffisantes pour établir que la modélisation conjointe proposée est une alternative très compétitive dans la prédiction et la qualité d'ajustement.

Le modèle proposé est sensible aux modifications. Par exemple, le modèle peut être facilement adapté pour inclure les effets fixes associés à un ensemble de covariables $\{x_{i,i',j,p}\}$, avec le prédicteur linéaire $x_{i,i',j}^T\beta_j - \|u_{i,j} - u_{i',j}\|$, où $x_{i,i',j}^T\beta_j = \sum_{p=1}^{P} x_{i,i',j,p}\beta_{j,p}$ représente les modèles des données associées aux covariables. De plus, au lieu de considérer un modèle de distance, il est également possible de faire une formule basée sur d'autres effets latents, tels qu'un modèle factoriel (Hoff, 2009). Aussi, le modèle proposé peut être étendu pour gérer les réseaux orientés au moyen d'une différenciation des positions latentes en fonction des caractéristiques latentes de la « sortie », $u_{i,j}$ et des caractéristiques latentes de « l'arrivée », $v_{i,j}$, de sorte que les effets latents peuvent être exprimés comme $-\|u_{i,j} - v_{i',j}\|$. Enfin, dans le même esprit que Green et Hastie (2009), il est également possible de faire une spécification trans-dimensionnelle du modèle où on affecte une distribution a priori à la dimension latente K.

Comme observé dans l'analyse des réseaux de la salle de câblage dans la Section 7, les positions sociales sont susceptibles d'afficher des modèles de regroupement, qui peuvent être modélisés directement en incluant des paramètres d'attribution aux groupes $\xi_1, \ldots, \xi_I$ dans la distribution a priori. Plus précisément, on peut supposer que les acteurs du système sont regroupés en H groupes, chacun occupant une position θ_h dans l'espace social. Ainsi, chacun des acteurs se voit attribuer une position globale η_i correspondant à un écart normal de la position du groupe auquel elle appartient, de sorte que $\eta_i | \theta_h, \kappa^2, \xi_i \overset{ind}{\sim} N(\theta_{\xi i}, \kappa^2 \mathrm{I})$, où $\xi_i = h$ signifie que l'acteur i fait partie du groupe h.

Une grande partie de la littérature sur les critères d'information s'est concentrée sur les critères d'information bayésien (BIC). Cependant, il a été démontré que le BIC est inapproprié pour les modèles hiérarchiques, car la structure hiérarchique implique que le nombre effectif de paramètres est statistiquement inférieur au nombre réel de paramètres. D'un autre côté, dans le contexte de ce travail, le BIC n'est pas jugé utile car il diffère notoirement des critères d'information de Watanabe-Akaike (WAIC). Dans le WAIC, les critères d'information ne sont pas motivés par une estimation de l'ajustement prédictif mais par l'objectif d'approcher la densité de probabilité marginale des données. Pour cette raison, il est tout à fait possible que le modèle ait une bonne prédiction et un WAIC faible, mais, en raison de la fonction de pénalité excessive du BIC (voir Tableau 4), il peut avoir un BIC relativement élevé, c'est-à-dire pauvre.

Enfin, il est intéressant d’étudier des implémentations du modèle qui permettent d'étudier les "grands" réseaux. Il s'agit actuellement d'un domaine de recherche actif. Des travaux à ce sujet ont été réalisés par Ma et Ma, 2017 ; Spencer et al., 2020 ainsi qu’Aliverti et Russo, 2022.

Références

Albert, J. H., & Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. In Journal of the American Statistical Association (Vol. 88, Issue 422, pp. 669–679). Informa UK Limited.https://doi.org/10.1080/01621459.1993.10476321

Aldous, D. J. (1985). Exchangeability and related topics. In Lecture Notes in Mathematics (pp. 1–198). Springer Berlin Heidelberg. https://doi.org/10.1007/bfb0099421

Aliverti, E., & Russo, M. (2022). Stratified Stochastic Variational Inference for High-Dimensional Network Factor Model. In Journal of Computational and Graphical Statistics (Vol. 31, Issue 2, pp. 502–511). Informa UK Limited. https://doi.org/10.1080/10618600.2021.1984929

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The Diffusion of Microfinance. In Science (Vol. 341, Issue 6144). American Association for the Advancement of Science (AAAS). https://doi.org/10.1126/science.1236498

Borg, I. & Groenen, P. J. (2005). In Springer Series in Statistics. Springer New York. https://doi.org/10.1007/0-387-28981-x

D'Angelo, S., Alfò, M., & Fop, M. (2020). *Model-based Clustering for Multivariate Networks. arXiv: Methodology*.

D'Angelo, S., Murphy, T.B., & Alfò, M. (2019). *Latent space modelling of multidimensional networks with application to the exchange of votes in Eurovision song contest. The Annals of Applied Statistics.*

D'Angelo, Silvia & Alfo, Marco & Murphy, Thomas. (2018). Node-specific effects in latent space modelling of multidimensional networks.

Durante, D., & Dunson, D. B. (2014). Nonparametric Bayes dynamic modelling of relational data. In Biometrika (Vol. 101, Issue 4, pp. 883–898). Oxford University Press (OUP). https://doi.org/10.1093/biomet/asu040

Durante, D., & Dunson, D. B. (2014). Bayesian Inference and Testing of Group Differences in Brain Networks. ArXiv. https://doi.org/10.48550/ARXIV.1411.6506

Gamerman, D., & Lopes, H. F. (2006). Markov chain Monte Carlo: Stochastic simulation for bayesian inference. Chapman & Hall/CRC.

Gao, L. L., Witten, D., & Bien, J. (2019). Testing for Association in Multi-View Network Data (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1909.11640

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian Data Analysis. Chapman and Hall/CRC.

Gelman, A., Hwang, J., & Vehtari, A. (2013). Understanding predictive information criteria for Bayesian models. In Statistics and Computing (Vol. 24, Issue 6, pp. 997–1016). Springer Science and Business Media LLC. https://doi.org/10.1007/s11222-013-9416-2

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. In Statistical Science (Vol. 7, Issue 4). Institute of Mathematical Statistics. https://doi.org/10.1214/ss/1177011136

Gollini, I., & Murphy, T. B. (2016). Joint Modeling of Multiple Network Views. In Journal of Computational and Graphical Statistics (Vol. 25, Issue 1, pp. 246–265). Informa UK Limited. https://doi.org/10.1080/10618600.2014.978006

Green, P. J. and Hastie, D. I. (2009). Reversible jump mcmc.Genetics, 155(3):1391–1403.

Gupta, S., Sharma, G., & Dukkipati, A. (2018). *Evolving Latent Space Model for Dynamic Networks. ArXiv, abs/1802.03725.*

Haario, H., Saksman, E., & Tamminen, J. (2001). An Adaptive Metropolis Algorithm. In Bernoulli (Vol. 7, Issue 2, p. 223). JSTOR. https://doi.org/10.2307/3318737

Han, Q., Xu, K. S., & Airoldi, E. M. (2014). Consistent estimation of dynamic and multi-layer block models. ArXiv. https://doi.org/10.48550/ARXIV.1410.8597

Handcock, M. S., Raftery, A. E., & Tantrum, J. M. (2007). Model-based clustering for social networks. In Journal of the Royal Statistical Society: Series A (Statistics in Society) (Vol. 170, Issue 2, pp. 301–354). Wiley. https://doi.org/10.1111/j.1467-985x.2007.00471.x

Hoff, P.D. (2007). *Modeling homophily and stochastic equivalence in symmetric relational data. NIPS*, pages 657–664.

Hoff, P. D. (2005). Bilinear Mixed-Effects Models for Dyadic Data. In Journal of the American Statistical Association (Vol. 100, Issue 469, pp. 286–295). Informa UK Limited. https://doi.org/10.1198/016214504000001015

Hoff, P. D. (2009). A First Course in Bayesian Statistical Methods. In Springer Texts in Statistics. Springer New York. https://doi.org/10.1007/978-0-387-92407-6

Hoff, P. D. (2014). Multilinear tensor regression for longitudinal relational data. ArXiv. https://doi.org/10.48550/ARXIV.1412.0048

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent Space Approaches to Social Network Analysis. In Journal of the American Statistical Association (Vol. 97, Issue 460, pp. 1090–1098). Informa UK Limited. https://doi.org/10.1198/016214502388618906

Hoover, D. N. (1982). Row-column exchangeability and a generalized model for probability.Exchangeability in probability and statistics (Rome, 1981), pages 281–291.

Kim, B., Lee, K. H., Xue, L., & Niu, X. (2018). A review of dynamic network models with latent variables. In Statistics Surveys (Vol. 12, Issue none). Institute of Mathematical Statistics. https://doi.org/10.1214/18-ss121

Kolaczyk, E. D., & Csárdi, G. (2014). Statistical Analysis of Network Data with R. In Use R! Springer New York, volume 65. https://doi.org/10.1007/978-1-4939-0983-4

Krackhardt, D. (1987). Cognitive social structures. In Social Networks (Vol. 9, Issue 2, pp. 109–134). Elsevier BV. https://doi.org/10.1016/0378-8733(87)90009-8

Krivitsky, P. N., & Handcock, M. S. (2008). Fitting Position Latent Cluster Models for Social Networks withlatentnet. In Journal of Statistical Software (Vol. 24, Issue 5). Foundation for Open Access Statistic. https://doi.org/10.18637/jss.v024.i05

Krivitsky, P. N., Handcock, M. S., Raftery, A. E., & Hoff, P. D. (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. In Social Networks (Vol. 31, Issue 3, pp. 204–213). Elsevier BV. https://doi.org/10.1016/j.socnet.2009.04.001

Li, W.J., Yeung, D.Y., & Zhihua, Z. (2011). Generalized latent factor models for social network analysis. In Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two (IJCAI'11). AAAI Press, 1705–1710.

Linkletter, C. D. (2007). Spatial process models for social network analysis (Doctoral dissertation, Simon Fraser University).

Ma, Z., & Ma, Z. (2017). Exploration of Large Networks with Covariates via Fast and Universal Latent Space Model Fitting (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.02372

Minhas, S., Hoff, P.D., & Ward, M.D. (2018). *Inferential Approaches for Network Analysis: AMEN for Latent Factor Models. Political Analysis, 27, 208 - 222.*

Nowicki, K., & Snijders, T. A. B. (2001). Estimation and Prediction for Stochastic Blockstructures. In Journal of the American Statistical Association (Vol. 96, Issue 455, pp. 1077–1087). Informa UK Limited. https://doi.org/10.1198/016214501753208735

Amini, A. A., Paez, M. S., & Lin, L. (2019). Hierarchical Stochastic Block Model for Community Detection in Multiplex Networks (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1904.05330

Paul, S., & Chen, Y. (2016). *Consistent community detection in multi-relational data through restricted multi-layer stochastic blockmodel. Electronic Journal of Statistics, 10, 3807-3870.*

Paul, S., & Chen, Y. (2020). Spectral and matrix factorization methods for consistent community detection in multi-layer networks. *Annals of Statistics*, *48*(1), 230-250. https://doi.org/10.1214/18-AOS1800

Polson, N. G., Scott, J. G., & Windle, J. (2013). Bayesian Inference for Logistic Models Using Pólya–Gamma Latent Variables. In Journal of the American Statistical Association (Vol. 108, Issue 504, pp. 1339–1349). Informa UK Limited. https://doi.org/10.1080/01621459.2013.829001

Raftery, A. E., Niu, X., Hoff, P. D., & Yeung, K. Y. (2012). Fast Inference for the Latent Space Network Model Using a Case-Control Approximate Likelihood. In Journal of Computational and Graphical Statistics (Vol. 21, Issue 4, pp. 901–919). Informa UK Limited. https://doi.org/10.1080/10618600.2012.679240

Reyes, P., & Rodriguez, A. (2016). Stochastic blockmodels for exchangeable collections of networks (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1606.05277

Dickson, W. J., & Roethlisberger, F. J. (2004). Management and the Worker. Routledge. https://doi.org/10.4324/9780203503010

Salter-Townshend, M., & McCormick, T. H. (2017). LATENT SPACE MODELS FOR MULTIVIEW NETWORK DATA. The Annals of Applied Statistics, 11(3), 1217–1244. http://www.jstor.org/stable/26362225

Schweinberger, M., & Snijders, T. A. B. (2003). 10. Settings in Social Networks: A Measurement Model. In Sociological Methodology (Vol. 33, Issue 1, pp. 307–341). SAGE Publications. https://doi.org/10.1111/j.0081-1750.2003.00134.x

Sewell, D. K., & Chen, Y. (2015). Latent Space Models for Dynamic Networks. In Journal of the American Statistical Association (Vol. 110, Issue 512, pp. 1646–1657). Informa UK Limited. https://doi.org/10.1080/01621459.2014.988214

Sewell, D. K., & Chen, Y. (2016). Latent space models for dynamic networks with weighted edges. In Social Networks (Vol. 44, pp. 105–116). Elsevier BV. https://doi.org/10.1016/j.socnet.2015.07.005

Sewell, D. K., & Chen, Y. (2017). Latent Space Approaches to Community Detection in Dynamic Networks. ArXiv. https://doi.org/10.48550/ARXIV.2005.08276

Sewell, D. K., & Chen, Y. (2020). Latent Space Models for Dynamic Networks. ArXiv. https://doi.org/10.48550/ARXIV.2005.08808

Sosa, J. (2017). A Latent Space Approach for Cognitive Social Structures Modeling and Graphical Record Linkage. *UC Santa Cruz*. ProQuest ID: Sosa_ucsc_0036E_11428. Merritt ID: ark:/13030/m55x75wm.

Spencer, N. A., Junker, B., & Sweet, T. M. (2020). Faster MCMC for Gaussian Latent Position Network Models (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2006.07687

Swartz, T. B., Gill, P. S., & Muthukumarana, S. (2015). A Bayesian approach for the analysis of triadic data in cognitive social structures. In Journal of the Royal Statistical Society: Series C (Applied Statistics) (Vol. 64, Issue 4, pp. 593–610). Wiley. https://doi.org/10.1111/rssc.12096

Turnbull, K. (2020).Advancements in latent space network modelling. PhD thesis, Lancaster University.

Wang, L., Zhang, Z., & Dunson, D. (2017). Common and Individual Structure of Brain Networks (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1707.06360

Watanabe, S. (2010). *Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory*. *ArXiv, abs/1004.2316*.

A. Algorithme de MCMC pour le modèle de réseaux multiples

L'algorithme de MCMC itère sur les paramètres du modèle $\breve{Y}$ donnés dans la Section 5.1. Dans la mesure du possible, les distributions conditionnelles complètes sont échantillonnées comme dans un échantillonneur de Gibbs habituel. Sinon, des versions adaptatives avec des étapes de Metropolis-Hastings sont utilisées (par exemple, Haario et al., 2001). Alternativement, dans le même esprit qu'Albert et Chib (1993), des variables aléatoires suivant une distribution Polya-Gamma peuvent être introduites pour faciliter le calcul (Polson et al., 2013).

Dans ce cas, la distribution postérieure conjointe est donnée par :

$$p(\breve{Y}|Y)=\prod_{j,i<i'} Ber(y_{i,i',j}|\theta_{i,i',j})*\prod_{i,j} N(u_{i,j}|\eta_i,\sigma_j^2\mathrm{I})*\prod_j N(\zeta_j|\mu_\zeta,\tau_\zeta^2)*\prod_j IGam(\sigma_j^2|\alpha_\sigma,\beta_\sigma)$$
$$*\prod_i N(\eta_i|\theta,\kappa^2\mathrm{I})*N(\theta|\mu_{0,}\Sigma_0)*IGam(\kappa^2|\alpha_\kappa,\beta_\kappa)*N(\mu_\zeta|\mu_0,\sigma_0^2)*IGam(\tau_\zeta^2|a_\zeta,b_\zeta)$$
$$*Unif(\alpha_\sigma|c_\sigma,d_\sigma)*Gam(\beta_\sigma|a_\sigma,b_\sigma)$$

où $\theta_{i,i'}=expit(\zeta_j-\|u_{i,j}-u_{i',j}\|)$ indique la probabilité d'une interaction entrez les acteurs i et i'. Pour un ensemble d'hyper-paramètres donné, μ_0, Σ_0, a_κ, b_κ, μ_0, σ_0^2, a_ζ, b_ζ, a_σ, b_σ, c_σ et d_σ, l'algorithme génère un nouvel état $\breve{Y}^{(b+1)}$ à partir de l'état actuel $\breve{Y}^{(b)}, b=1,\ldots,B$, de la façon suivante:

1. Déterminer $u_{i,j}^{(b+1)}, i=1,\ldots,I, j=1,\ldots,J,$ en suivant algorithme de Métropolis-Hastings, en considérant la distribution conditionnelle complète:

$$p(u_{i,j}|rest) \propto \prod Ber(y_{i,i',j}|expit(\zeta_j-\|u_{i,j}-u_{i,j'}\|))$$
$$*\prod Ber(y_{i',i,j}|expit(\zeta_j-\|u_{i,j}-u_{i,j'}\|))*N(u_{i,j}|\eta_i,\sigma_j^2\mathrm{I}).$$

2. Déterminer $\zeta_j^{(b+1)}=1,\ldots,J$ en suivant un algorithme de Metropolis-Hastings, en considérant la distribution conditionnelle complète :

$$p(\zeta_j|rest) \propto \prod Ber(expit(\zeta_j-\|u_{i,j}-u_{i',j}\|))*N(\zeta_j|\mu_\zeta,\tau_\zeta^2)$$

3. Déterminer $(\sigma_i^2)^{(b+1)}, j=1,\ldots,J$ d'une distribution inverse-gamma

$$(\sigma_j^2)|rest \sim IGam(\alpha_\sigma+\frac{IK}{2},\beta_\sigma+\frac{1}{2}\sum_i (u_{i,j}-\eta_i)^T(u_{i,j}-\eta_i))$$

4. Déterminer $\eta_i^{(b+1)}, i=1,\ldots,I$ d'une distribution normale multivariée

$$\eta_i|rest \sim N([\frac{1}{\kappa^2}+\sum_j \frac{1}{\sigma_j^2}]^{-1}[\frac{1}{\kappa^2}\theta+\sum_j \frac{1}{\sigma_j^2}u_{i,j}],[\frac{1}{\kappa^2}+\sum_j \frac{1}{\sigma_j^2}]^{-1}\mathrm{I})$$

5. Déterminer θ d'une distribution normale multivariée

$$\theta|rest \sim N\left([\Sigma_0^{-1} + \frac{I}{\kappa^2}\mathrm{I}]^{-1}[\Sigma_0^{-1}\mu_0 + \frac{1}{\kappa^2}\sum_i \eta_i], [\Sigma_0^{-1} + \frac{I}{\kappa^2}\mathrm{I}]^{-1}\right)$$

6. Déterminer κ^2 d'une distribution inverse-gamma

$$\kappa_h^2|rest \sim IGam\left(\alpha_\kappa + \frac{IK}{2}, \beta_\kappa + \frac{1}{2}\sum_i (\eta_i - \theta)^T(\eta_i - \theta)\right)$$

7. Déterminer μ_ζ d'une distribution normale

$$\mu_\zeta|rest \sim N\left([\frac{1}{\sigma_0^2} + \frac{J}{\tau_\zeta^2}]^{-1}[\frac{\mu_0}{\sigma_0^2} + \frac{\sum_j \zeta_j}{\tau_\zeta^2}], [\frac{1}{\sigma_0^2} + \frac{J}{\tau_\zeta^2}]^{-1}\right)$$

8. Détermine τ_ζ^2 d'une distribution inverse-gamma

$$\tau_\zeta^2|rest \sim IGam\left(a_\zeta + \frac{J}{2}, b_\zeta + \frac{1}{2}\sum_j (\zeta_j - \mu_\zeta)^2\right)$$

9. Déterminer α_σ en suivant l'algorithme de Metropolis-Hastings, en considérant la distribution conditionnelle complète

$$\log p(\alpha_\sigma|rest) \propto J[\alpha_\sigma \log \beta_\sigma - \log \Gamma(\alpha_\sigma)] - \alpha_\sigma \sum_j \log \sigma_j^2$$

10. Déterminer β_σ d'une distribution Gamma

$$\beta_\sigma|rest \sim Gam\left(a_\sigma + J\alpha_\sigma, b_\sigma\right) + \sum_j \frac{1}{\sigma_j^2}$$

B. Notation

La valeur absolue d'un nombre réel x est notée $|x|$, et la fonction Gamma avec $\Gamma(x)=\int_0^\infty u^{x-1}e^{-u}du$. 0 et 1 sont utilisés pour désigner les vecteurs colonnes dont les entrées sont toutes égales à 0 et 1, respectivement, et I pour désigner la matrice d'identité. La transposée d'un vecteur x est notée x^T ; de même pour les matrices. Aussi, si X est une matrice carrée, on utilise $tr(X)$ pour désigner sa trace et X^{-1} pour faire référence à son inverse. La norme d'un vecteur x, donnée par $\sqrt{(x^T x)}$, est notée $\|x\|$. D'autre part, $p(\cdot|\cdot)$ est utilisé pour désigner une fonction de densité de probabilité conditionnelle ; de même avec $p(\cdot)$, qui indique une distribution de probabilité marginale. La même notation est utilisée pour les fonctions de densité continues et les fonctions de masse de probabilité discrètes. Bien qu'il s'agisse d'un abus de notation, la distribution conditionnelle complète d'un paramètre θ compte tenu du reste des paramètres et des données est notée $p(\theta|rest)$.