

Modèle à Blocs Stochastiques corrigé en degrés pour des graphes dynamiques

Paul Riverain^{*,**} Simon Fossier^{**} Mohamed Nadif^{*}

^{*}Université de Paris, LIPADE, 75006 Paris, France <prénom.nom@u-paris.fr>

^{**}Thales <prénom.nom@thalesgroup.com>

Résumé. Le Modèle à Blocs Stochastiques (SBM) permet d’offrir des méthodes de classification non supervisée des nœuds d’un graphe. Dans ce papier, nous proposons une extension de ce modèle pour des graphes dynamiques à temps discret permettant de modéliser une plus grande variété de graphes. Le modèle proposé génère des graphes possédant une structure en classes dynamiques ainsi que des paramètres de correction des degrés dépendant du temps qui permettent de modéliser des graphes dont les nœuds ont des degrés variables ou dont le nombre d’arêtes total varie de façon importante dans le temps. Pour l’inférence, nous proposons un algorithme basé sur une variante variationnelle de l’algorithme EM et un lissage temporel des paramètres, afin de limiter le problème de *label-switching* local. L’application de l’algorithme sur des données réelles a fourni des résultats qui illustrent l’intérêt de notre approche.

1 Introduction

Le Modèle à Blocs Stochastiques (SBM) (Snijders et Nowicki, 1997; Daudin et al., 2008) est un modèle statistique pour les données relationnelles qui fournit un modèle parcimonieux, grâce à l’implication croisée de variables latentes, pour la classification non-supervisée des nœuds d’un graphe. Les réseaux réels dépendant souvent du temps (réseaux sociaux, réseaux de transport, réseaux de citations, . . .), l’extension du SBM pour des données dynamiques est d’intérêt pour l’étude de ces réseaux. Dans un contexte de *clustering* de graphe à temps discret, le but du clustering est de trouver une structure en classes à chaque instant du graphe, tout en maintenant une cohérence dans la dynamique des classes (Fu et al., 2009; Xu et Hero, 2014; Sewell et Chen, 2016; Matias et Miele, 2017). Un SBM statique appliqué à chaque instant du graphe serait ici inapproprié car il ne saisirait pas la dynamique du réseau.

Dans ce papier, nous nous plaçons dans le cadre de multigraphes dynamiques à temps discret, que l’on peut voir aussi comme des graphes dynamiques pondérés par des valeurs entières, et nous cherchons à prendre en compte deux spécificités de ces réseaux dynamiques qui seront décrites ci-après.

Tout d’abord, dans ces réseaux, le nombre d’interactions à chaque instant (le nombre d’arêtes dans chaque multigraphe) peut évoluer de façon importante dans le temps, et certains nœuds peuvent entrer et quitter le réseau à certains instants. Par exemple, dans un réseau de transports, le nombre total de voyages varie beaucoup entre l’ouverture du réseau et l’heure de pointe.

Ensuite, les réseaux réels présentent une grande variabilité dans les degrés des nœuds, suivant parfois une loi de puissance (Barabási et Albert, 1999). Cependant, le SBM classique tend à mettre dans la même classe des nœuds de degré similaire (Karrer et Newman, 2011). Nous nous intéressons donc à ces variantes corrigées en degré, qui considèrent que des nœuds d'une même classe partagent un même *profil de connectivité*, c'est à dire qu'ils se connectent dans les mêmes proportions aux nœuds des autres classes.

Notations Les vecteurs, matrices et tenseurs sont notés en gras. Nous considérons un graphe dynamique à temps discret à N nœuds, représenté par une suite de matrices d'adjacence $\mathbf{X} = (\mathbf{X}^t)_{t \in \{1, \dots, T\}}$, où T désigne le nombre d'instantants durant lesquels on considère le graphe. \mathbf{X}^t est la matrice d'adjacence pondérée du graphe orienté; $X_{ij}^t \in \mathbb{N}$ et $X_{ii}^t = 0$. Soit $(Z_i^t)_{t \in \{1, \dots, T\}}$ le processus latent associé au nœud i , à valeurs dans $\{1, \dots, K\}$, où K est le nombre de *clusters*. Si le nœud i appartient au cluster k à l'instant t , on note $Z_i^t = k$ ou de façon équivalente $Z_{ik}^t = 1$. Soit $\mathcal{P}(\lambda)$ une distribution de Poisson de paramètre λ et $\phi(\cdot; \lambda)$ sa fonction de masse. \mathbb{H} correspond à l'opérateur d'entropie et $\mathbb{1}$ est la fonction indicatrice. Les sommes et produits correspondant respectivement aux instantants, nœuds et clusters seront indicés avec les lettres t, i, j, k and ℓ avec des limites de variation implicites. Ainsi $\sum_t, \sum_i, \sum_j, \sum_k$ and \sum_ℓ correspondent respectivement à $\sum_{t=1}^T, \sum_{i=1}^N, \sum_{j=1}^N, \sum_{k=1}^K$ and $\sum_{\ell=1}^K$.

2 Travaux liés

Dans le but de modéliser une grande variété de réseaux réels, nous nous intéresserons tout d'abord aux variantes du SBM qui permettent de prendre en compte la variabilité des degrés des nœuds, puis nous présenterons certaines extensions du SBM dans le cadre dynamique.

Correction des degrés dans SBM Considérons tout d'abord ce modèle général de graphes statiques : $X_{ij}|Z_{ik}Z_{j\ell} = 1 \sim \mathcal{F}(\lambda_{ijk\ell})$, où les arêtes X_{ij} sont générées par une distribution paramétrique \mathcal{F} . Karrer et Newman (2011) proposent de prendre en compte la variabilité des degrés des nœuds pour des graphes non-orientés en considérant $\lambda_{ijk\ell} = \mu_i \mu_j \gamma_{k\ell}$ pour une distribution de Poisson. La normalisation $\sum_i \mu_i Z_{ik} = 1$, pour chaque cluster k , est proposée afin de rendre le modèle identifiable. La correction μ_i peut alors être interprétée comme étant la probabilité qu'une arête du cluster du nœud i soit connectée à i . Cependant, la contrainte sur la variable latente Z rend son utilisation difficile dans le cadre de l'algorithme EM et requiert alors l'utilisation de méthodes heuristiques pour l'inférence. Qiao et al. (2017) développent cette idée, pour modéliser des réseaux invariants à l'échelle, en utilisant des arêtes binaires et $\lambda_{ijk\ell} = \gamma_{k\ell}^{1+\mu_i+\mu_j}$. Les auteurs montrent alors, en faisant l'hypothèse que le degré d'un nœud est essentiellement dû à son cluster (homophilie), que le degré des nœuds converge vers une variable aléatoire qui suit une loi de puissance lorsque le nombre de nœuds tend vers l'infini. Cependant, l'inférence dans ce modèle est difficile et ne donne pas de formules analytiques. Dans le contexte du modèle des *blocs latents* (LBM), et pour des arêtes distribuées selon une loi de Poisson, Ailem et al. (2017) proposent de considérer une correction de degré pour chaque colonne et chaque ligne de la matrice d'adjacence du graphe bipartite $\lambda_{ijk\ell} = \mu_i \nu_j \gamma_{k\ell}$. Les paramètres de correction sont normalisés de façon à ce que $\mathbb{E}(\sum_j X_{ij}) = \mu_i$, ce qui permet de ne pas avoir à estimer de correction de degré dans un contexte de classification.

Cette normalisation n'est cependant pas applicable dans le contexte du SBM, soit à cause des termes diagonaux de la matrice d'adjacence (si on considère ceux-ci non nuls), soit à cause d'une structure de dépendance plus contraignante pour le SBM que le LBM.

Variantes dynamiques du SBM Matias et Miele (2017) présente un modèle dans lequel un processus latent à valeurs discrètes est associé à chaque nœud du graphe et la matrice de connectivité du graphe évolue dans le temps. Ces processus correspondent à des chaînes de Markov indépendantes et identiquement distribuées. Les auteurs montrent que, pour des raisons d'identifiabilité, soit les clusters peuvent avoir une matrice de connectivité variable dans le temps et doivent être définis par les nœuds qu'ils contiennent – et sont donc constants –, soit les nœuds peuvent changer de cluster au cours du temps, mais leur matrice de connectivité doit être contrainte de manière à imposer une caractéristique constante à chaque cluster ($\gamma_{kk}^t = \gamma_{kk}$). Pour l'inférence, les auteurs proposent un algorithme – `dynsbm`¹ – qui repose sur une variante variationnelle de EM. Rastelli et al. (2018) proposent un modèle similaire dans un contexte bayésien, mais dans lequel la matrice de connectivité des clusters est constante. Afin d'offrir suffisamment de flexibilité au modèle, il est alors nécessaire de considérer un grand nombre de clusters, dont la plupart seront vides à certains instants. L'inférence est réalisée de manière similaire à celle de Corneli et al. (2016). Ainsi, dans les deux précédents modèles, les connectivités intra-classe γ_{kk}^t sont constantes dans le temps, ce qui impose de considérer un grand nombre de clusters si l'on souhaite modéliser un graphe dont le nombre d'arêtes dans toutes les classes évolue beaucoup au cours du temps.

3 Modèle proposé

3.1 Définition du modèle

Soient $(Z_i^t)_i$, N chaînes de Markov indépendantes à valeurs dans $\{1, \dots, K\}$, de distribution initiale multinomiale de paramètre α , où $\forall i, \alpha_k = P(Z_i^1 = k)$ et une matrice de transition π , telle que $\forall i, \forall t \in \{2, \dots, T\}, \pi_{k\ell} = P(Z_i^t = \ell | Z_i^{t-1} = k)$. Les poids des arêtes orientées X_{ij}^t du graphe sont générés selon une loi de Poisson dont l'intensité est déterminée par trois termes : la marge entrante μ_i^t du nœud de départ i à l'instant t , la marge sortante ν_j^t du nœud d'arrivée j , ainsi que le terme de connectivité $\gamma_{Z_i^t Z_j^t}$ correspondant aux interactions entre le cluster du nœud i et celui du nœud j . Ainsi, on écrit :

$$X_{ij}^t | Z_{ik}^t Z_{j\ell}^t = 1 \sim \mathcal{P}(\mu_i^t \nu_j^t \gamma_{k\ell}). \quad (1)$$

3.2 Inférence

Nous cherchons à inférer de façon jointe les variables latentes \mathbf{Z} et les paramètres du modèle $\theta = \{\alpha, \pi, \mu, \nu, \gamma\}$. La log-vraisemblance des données complètes est donnée par :

$$\log P(\mathbf{X}, \mathbf{Z}; \theta) = \sum_{ik} Z_{ik}^1 \log \alpha_k + \sum_{t \geq 2} \sum_{ik\ell} Z_{ik}^{t-1} Z_{i\ell}^t \log \pi_{k\ell} + \sum_t \sum_{i,j|i \neq j} \sum_{k\ell} Z_{ik}^t Z_{j\ell}^t \log \phi(X_{ij}^t; \mu_i^t \nu_j^t \gamma_{k\ell}). \quad (2)$$

1. <https://cran.r-project.org/web/packages/dynsbm/dynsbm.pdf>

Inférence variationnelle Étant donné qu'il n'est pas possible de maximiser directement la log-vraisemblance du modèle avec l'algorithme EM, nous employons une approximation variationnelle de la distribution postérieure décrite dans (Govaert et Nadif, 2005; Daudin et al., 2008). Nous utilisons une distribution variationnelle Q , parmi les distributions sur l'espace latent qui se factorisent comme une chaîne de Markov : $Q(\mathbf{Z}) = \prod_i Q(Z_i^1) \prod_{t \geq 2} Q(Z_i^t | Z_i^{t-1})$, de façon similaire à Matias et Miele (2017), et nous définissons $q(i, k) = Q(Z_i^1 = k)$, $q(t, i, k, \ell) = Q(Z_i^t = \ell | Z_i^{t-1} = k)$. On peut alors calculer les probabilités marginales de façon récursive : $q(t, i, k) = Q(Z_i^t = k) = \sum_{k'} q(t-1, i, k')q(t, i, k', k)$. Les paramètres de la distribution variationnelle et du modèle sont obtenus en maximisant $F(\mathbf{q}, \boldsymbol{\theta})$, une borne inférieure de la log-vraisemblance du modèle :

$$\log P(\mathbf{X}, \mathbf{Z}; \boldsymbol{\theta}) \geq F(\mathbf{q}, \boldsymbol{\theta}) = \mathbb{E}_Q(\log P(\mathbf{X}, \mathbf{Z}; \boldsymbol{\theta})) + \mathbb{H}(Q). \quad (3)$$

Étapes E et M Pour l'étape E, nous utilisons l'approche heuristique proposée par Matias et Miele (2017), en mettant à jour les probabilités variationnelles de transition avec (??, 4, ??) avec les notations $\phi_{ijk\ell}^t = \phi(X_{ij}^t; \mu_i^t \nu_j^t \gamma_{k\ell})$ et $\Phi_{ijk\ell}^t = \phi_{ijk\ell}^t \phi_{j\ell k}^t$. Ces formules sont obtenues en remarquant que chaque nœud peut être considéré de façon indépendante. Nous appliquons alors une unique itération d'un algorithme de montée par coordonnées avec des mises à jour de Jacobi pour un instant t (donc réalisées en parallèle) en considérant les probabilités variationnelles fixées aux autres instants.

$$q(t, i, k, \ell) \propto \pi_{k\ell} \prod_{j \neq i} \prod_{\ell'} \Phi_{ij\ell\ell'}^t q^{(t, j, \ell')} \quad (4)$$

Afin de mettre à jour les paramètres du modèle dans l'étape M, on augmente $F(\mathbf{q}, \boldsymbol{\theta})$ par rapport à $\boldsymbol{\theta}$ en maximisant F pour chaque paramètre, conditionnellement aux autres. On met tout d'abord à jour les paramètres de mélange α , π and γ car ils ne dépendent que de \mathbf{q} , puis on met à jour $\boldsymbol{\mu}$, et finalement $\boldsymbol{\nu}$.

$$\widehat{\alpha}_k \propto \sum_i q(i, k), \quad \widehat{\pi}_{k\ell} \propto \sum_{t \geq 2} \sum_i q(t-1, i, k)q(t, i, k, \ell), \quad (5)$$

$$\widehat{\mu}_i^t = \frac{\sum_{j \neq i} \sum_{k\ell} q(t, i, k)q(t, j, \ell)X_{ij}^t}{\sum_{j \neq i} \sum_{k\ell} q(t, i, k)q(t, j, \ell)\nu_j^t \gamma_{k\ell}}, \quad \widehat{\nu}_j^t = \frac{\sum_{i \neq j} \sum_{k\ell} q(t, i, k)q(t, j, \ell)X_{ij}^t}{\sum_{i \neq j} \sum_{k\ell} q(t, i, k)q(t, j, \ell)\mu_i^t \gamma_{k\ell}}, \quad (6)$$

$$\widehat{\gamma}_{k\ell} = \frac{\sum_t \sum_{i, j | i \neq j} q(t, i, k)q(t, j, \ell)X_{ij}^t}{\sum_t \sum_{i, j | i \neq j} q(t, i, k)q(t, j, \ell)\mu_i^t \nu_j^t}. \quad (7)$$

Métriques globales et locales En clustering, nous faisons face au problème du *label-switching*, c'est à dire que les classes ne peuvent être retrouvées qu'à une permutation près. Dans le cadre du clustering dynamique, nous devons aussi prendre en compte le problème difficile du *label-switching* local, qui consiste à associer les classes trouvées à un instant donné avec celles de l'instant suivant. Afin d'évaluer la qualité des partitions trouvées, nous considérons l'indice de Rand ajusté (ARI) (Hubert et Arabie, 1985)). Nous utilisons deux autres versions de cette métrique : une métrique locale correspondant à la moyenne de l'ARI à chaque instant, sur T partitions de N points, ainsi qu'une métrique globale, correspondant à l'ARI d'une unique partition de TN points.

Lissage temporel des paramètres Lors de nos expériences avec le modèle et l’algorithme décrits ci-dessus, tout comme avec `dynsbm`, nous avons observé que, lorsque les classes sont peu stables dans le temps, il y a souvent une différence importante entre les métriques globales et locales. Ceci est révélateur d’un problème de *label-switching* local, dû au fait qu’à chaque itération de l’algorithme, le critère $F(\mathbf{q}, \boldsymbol{\theta})$ est optimisé en termes de paramètres de mélange et en termes de paramètre des densités conditionnelles, résultant potentiellement en une configuration où les paramètres expliquent mieux les classes à chaque instant, mais où leur dynamique est moins bien saisie. Il n’y a en effet pas de garantie pour l’algorithme que nous proposons, tout comme pour `dynsbm`, que le critère augmente à chaque itération.

Afin de réduire ce problème, nous proposons de garder les marges du modèle constantes durant une première phase de l’inférence, puis de relâcher progressivement cette contrainte au cours des itérations suivantes de EM. L’idée est de partir d’un modèle avec une capacité réduite afin de mettre l’accent sur l’estimation des paramètres de mélange α et π . Une fois que ceux-ci sont estimés dans le modèle à capacité réduite, on relâche la contrainte sur les marges afin de rendre leur capacité aux densités conditionnelles, et ceci de manière progressive, afin de ne pas provoquer de *label-switching* local. Pour cela, on peut observer que si μ_i^t est constant dans le temps, alors on l’estime de la façon suivante : $\sum_t n_{\mu_i}^t / \sum_t d_{\mu_i}^t$, où $n_{\mu_i}^t$ et $d_{\mu_i}^t$ sont respectivement les numérateurs et dénominateurs de (6). Pour relâcher la contrainte progressivement, nous proposons un filtrage temporel du numérateur et du dénominateur des marges estimées, dont on va faire varier la bande passante au cours des itérations. Pour un signal $(x_t)_t$, $S_{W^\tau}(x)_t = \sum_{t'} W_{tt'}^\tau x_{t'}$, où $W_{tt'}^\tau$ est un noyau gaussien centré à t et dont la bande passante $\tau \in [0, 1]$ contrôle le niveau de lissage (voir Fig. 1). On considère $n_{\mu_i}^t$ et $d_{\mu_i}^t$

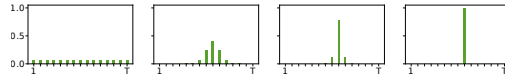


FIG. 1: Poids $W_{tt'}^\tau$ du lissage temporel en fonction de t' , pour $t = 8$, $\tau \in \{0, \frac{1}{3}, \frac{2}{3}, 1\}$

comme des signaux et on estime μ_i^t et ν_j^t avec (8). Le même principe peut être appliqué pour l’estimation de la matrice de connectivité présentée dans (Matias et Miele, 2017).

$$\tau \widehat{\mu}_i^t = \frac{S_{W^\tau}(n_{\mu_i}^t)_t}{S_{W^\tau}(d_{\mu_i}^t)_t}, \quad \tau \widehat{\nu}_j^t = \frac{S_{W^\tau}(n_{\nu_j}^t)_t}{S_{W^\tau}(d_{\nu_j}^t)_t}. \quad (8)$$

4 Résultats expérimentaux

Effet des marges Notons M_- le modèle sans marges ($X_{ij}^t | Z_{ik}^t Z_{j\ell}^t = 1 \sim \mathcal{P}(\gamma_{k\ell}^t)$, avec $\forall k, \gamma_{kk}^t = \gamma_{kk}$ comme décrit dans (Matias et Miele, 2017)) et M_+ le modèle avec marges ($X_{ij}^t | Z_{ik}^t Z_{j\ell}^t = 1 \sim \mathcal{P}(\mu_i^t \nu_j^t \gamma_{k\ell}^t)$). Comme illustré qualitativement Fig. 2, alors que M_- tend à regrouper à des nœuds de même degré, M_+ va regrouper des nœuds qui partagent un même *profil de connectivité*. De façon plus quantitative, nous avons pu observer que M_- ne peut distinguer des classes générées par M_+ . Cependant, le modèle M_+ obtient souvent des performances comparables à M_- lorsque les données sont générées par M_- et que les classes ne sont pas trop proches.

Modèle à Blocs Stochastiques dynamique avec correction de degrés

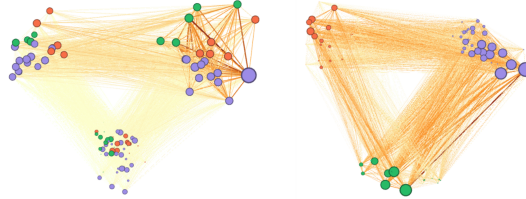


FIG. 2: Clusters obtenus respectivement avec M_- (gauche) et M_+ (droite). Le graphe initial est non-orienté et généré avec M_+ . La couleur des nœuds représente leur vraie classe et leur position dans le plan correspond aux clusters obtenus avec l'algorithme. Le poids des arêtes est représenté par leur couleur (jaune et rouge correspondent respectivement à un poids faible et élevé).

Trajets dans le réseau Transport For London Le jeu de données² compte tous les trajets effectués du 15 au 28 août 2019 entre les $N = 778$ stations du réseau de vélos libre-service de Londres. Le graphe dynamique est construit en agrégeant les trajets de chaque journée ($T = 14$). Le nombre de clusters ($K = 7$) est choisi à partir de la méthode du coude dans la log-vraisemblance des données complètes.

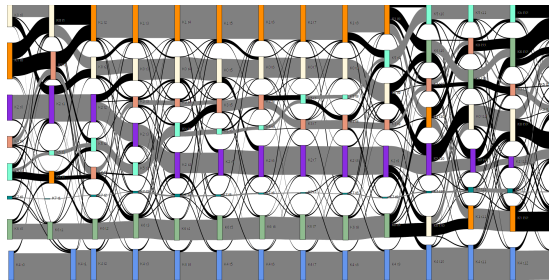


FIG. 3: Transitions de classes du 15 au 28 août 2019, représentées par un diagramme alluvial. Chaque cluster est représenté par une couleur, de taille proportionnelle au nombre de nœuds qu'il contient. Le temps est représenté de gauche à droite, et les transitions entre clusters entre instants consécutifs sont représentées par un flux coloré (gris pour des transitions intra-cluster, noir en inter-cluster). Le troisième cluster en partant du bas correspond aux de degré nul.

Les tailles des clusters ainsi que leurs transitions sont représentées sur le diagramme alluvial Fig. 3. On observe une dynamique des clusters assez stable du 16 au 23 août et de nombreuses transitions les 24 et 25 août. Ceci peut probablement être expliqué par la tenue du carnaval de Notting Hill (2,5 millions de participants) les 24 et 25 août. Les clusters obtenus le 18 août sont présentés Fig. 4a. On observe que les clusters obtenus sont très corrélés géographiquement, ce qui est cohérent pour des trajets à vélo. Les clusters obtenus sont aussi géographiquement plus étendus en périphérie qu'au centre, ce qui est cohérent. Sur Fig.4b sont représentés les clusters obtenus le même jour avec M_- . On observe que les clusters obtenus sont moins cohérents géographiquement que ceux obtenus avec M_+ . Fig. 4c et Fig. 4d pré-

2. <https://cycling.data.tfl.gov.uk/>

sentent respectivement une estimation de la densité des degrés des nœuds dans chaque cluster avec un noyau gaussien pour les modèles M_+ et M_- . On y observe que, dans le cas de M_- , la distribution des degrés est homogène dans chaque cluster, ce qui n'est pas le cas avec M_+ . Le modèle que nous proposons permet au contraire, grâce aux marges, de classifier un nœud indépendamment de son degré et de se concentrer sur son *profil de connectivité*.

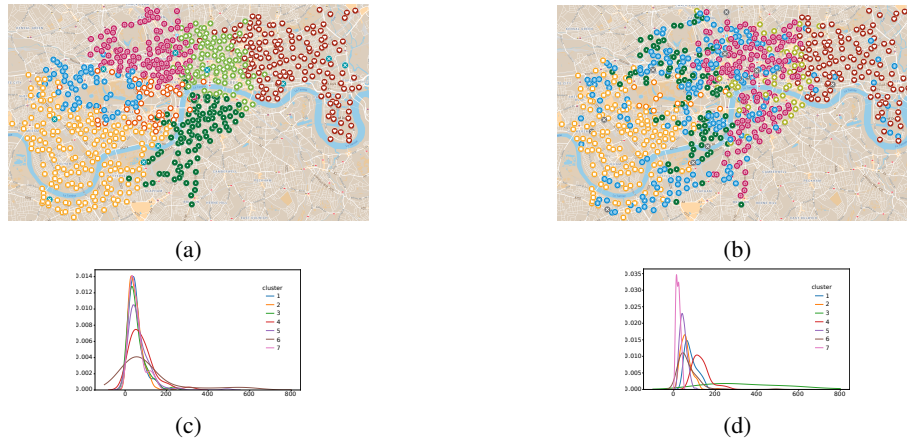


FIG. 4: Clusters obtenus pour le dimanche 18 août 2019 et estimation de la densité des degrés des nœuds dans chaque cluster avec un noyau gaussien *gauche* : M_+ , *droite* : M_-

5 Discussion et travaux futurs

Dans ce papier, nous avons présenté un modèle à Blocs Stochastiques dynamique avec correction de degrés, qui nous a permis de modéliser des graphes dynamiques possédant une importante variabilité dans les degrés des nœuds ainsi qu'un nombre d'arêtes qui évolue au cours du temps. Ceci n'était pas possible en maintenant un faible nombre de clusters avec les modèles proposés par Matias et Miele (2017) et Rastelli et al. (2018). Même si l'introduction de marges dans notre modèle est faite au prix de son identifiabilité, nous avons constaté que le fait de considérer une matrice de connectivité constante, en plus de permettre une interprétation plus aisée des clusters, contraignait suffisamment notre modèle pour obtenir une bonne stabilité dans la dynamique des clusters. Nous avons finalement proposé une technique d'inférence qui nous a permis d'estimer des paramètres dynamiques tout en réduisant les risques de *label-switching* local.

Références

- Ailem, M., F. Role, et M. Nadif (2017). Sparse Poisson Latent Block Model for Document Clustering. *IEEE Transactions on Knowledge and Data Engineering* 29(7), 1563–1576.
- Barabási, A.-L. et R. Albert (1999). Emergence of scaling in random networks. *Science* 286(5439), 509–512.

- Corneli, M., P. Latouche, et F. Rossi (2016). Exact ICL maximization in a non-stationary temporal extension of the stochastic block model for dynamic networks. *Neurocomputing* 192, 81 – 91.
- Daudin, J. J., F. Picard, et S. Robin (2008). A mixture model for random graphs. *Statistics and Computing* 18(2), 173–183.
- Fu, W., L. Song, et E. P. Xing (2009). Dynamic mixed membership blockmodel for evolving networks. In *ICML*, pp. 329–336.
- Govaert, G. et M. Nadif (2005). An EM Algorithm for the block mixture model. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(4).
- Hubert, L. et P. Arabie (1985). Comparing partitions. *Journal of Classification*.
- Karrer, B. et M. E. Newman (2011). Stochastic blockmodels and community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 83(1).
- Matias, C. et V. Miele (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. *Journal of the Royal Statistical Society. Series B : Statistical Methodology* 79(4), 1119–1141.
- Qiao, M., J. Yu, W. Bian, Q. Li, et D. Tao (2017). Improving stochastic block models by incorporating power-law degree characteristic. In *IJCAI International Joint Conference on Artificial Intelligence*, pp. 2620–2626.
- Rastelli, R., P. Latouche, et N. Friel (2018). Choosing the number of groups in a latent stochastic blockmodel for dynamic networks. *Network Science* 6(4), 469–493.
- Sewell, D. K. et Y. Chen (2016). Latent space models for dynamic networks with weighted edges. *Social Networks* 44, 105–116.
- Snijders, T. et K. Nowicki (1997). Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *Journal of Classification* 14, 75–100.
- Xu, K. S. et A. O. Hero (2014). Dynamic stochastic blockmodels for time-evolving social networks. *IEEE Journal on Selected Topics in Signal Processing* 8(4), 552–562.

Summary

Stochastic Block Models (SBM) provide a statistical tool for modeling and clustering network data. In this paper, we propose an extension of this model for discrete-time dynamic networks that takes into account the variability in node degrees, allowing us to model a broader class of networks. We develop a probabilistic model that generates temporal graphs with a dynamic cluster structure and time-dependent degree corrections for each node. Thanks to these degree corrections, the nodes can have variable degrees, allowing for more complex cluster structures and for model interactions that decrease or increase over time. The proposed model is compared to an existing model without degree correction and its advantages in terms of global performances are highlighted. We propose an inference procedure based on Variational EM that provides the means to estimate time-dependent parameters while reducing the risk of local label-switchings.