



Un modèle de sous-graphes aléatoires pour l'analyse de la dynamique d'un réseau historique

Charles Bouveyron, P Latouche, Rawya Zreik

► **To cite this version:**

Charles Bouveyron, P Latouche, Rawya Zreik. Un modèle de sous-graphes aléatoires pour l'analyse de la dynamique d'un réseau historique. 46èmes Journées de Statistique de la SFdS, Jun 2014, Rennes, France. pp.106-107. <hal-01199638>

HAL Id: hal-01199638

<https://hal.archives-ouvertes.fr/hal-01199638>

Submitted on 15 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UN MODÈLE DE SOUS-GRAPHE ALÉATOIRES POUR L'ANALYSE DE LA DYNAMIQUE D'UN RÉSEAU HISTORIQUE

C. Bouveyron¹, P. Latouche² & R. Zreik^{1,2}

¹ *Laboratoire MAP5, UMR CNRS 8145, Université Paris Descartes*

² *Laboratoire SAMM, EA 4543, Université Paris 1 Panthéon-Sorbonne*

Résumé. Ces dernières années, de nombreux modèles de graphes aléatoires ont été proposés pour extraire des informations à partir de réseaux. Le principe consiste à chercher des groupes de sommets ayant des profils de connexions homogènes. La plupart de ces modèles sont adaptés pour les réseaux statiques, c'est-à-dire ne prenant pas en compte la dimension temporelle, mais pouvant gérer différents types d'arêtes, qu'elles soient binaires ou discrètes. Les modèles existants doivent donc être adaptés pour être capables de gérer l'aspect temporel. Nous nous intéressons ici au modèle de sous-graphes aléatoires proposé par [1]. Ce modèle permet de modéliser les sous-graphes connus d'un réseau à l'aide de structures latentes cachées. Nous proposons d'intégrer un processus à espace d'états à ce modèle afin d'autoriser les classes, ainsi que la topologie du réseau, à évoluer au cours du temps. L'inférence est réalisée par l'intermédiaire d'un algorithme Bayés variationnel. Un critère de sélection de modèles est également dérivé. Nous appliquons cette méthodologie à un réseau décrivant les relations entre évêques en Gaule Mérovingienne.

Mots-clés. Réseau, sous-graphe, classification, algorithme VBEM.

1 Contexte de notre étude

Ce travail a été motivé par la nécessité d'analyser un réseau historique évoluant dans le temps, pour lequel une partition des sommets est donnée et dont les arêtes sont de type catégorielle. L'analyse porte sur les relations interindividuelles tissées dans le cadre des conciles réunis en Gaule au VI^e siècle. Le concile est une assemblée d'ecclésiastiques réunie sous l'autorité d'un évêque pour débattre des affaires ecclésiastiques. Pour le VI^e siècle, 46 conciles furent réunis en Gaule, qui est alors, au gré des partages successoraux, divisée entre plusieurs royaumes : Neustrie, Austrasie et Bourgogne, la Bourgogne et l'Aquitaine changeant plusieurs fois de mains au cours du VI^e siècle. La base de données regroupe 1331 personnages ayant occupé une ou plusieurs fonctions en Gaule entre 480 et 614 et les individus pour lesquels nous savons qu'ils ont eu des liens relationnels avec d'autres que ce soient des liens de parenté ou de connaissance. D'un point de vue historique, l'analyse et la comparaison de la structure organisationnelle de chacun des royaumes ou provinces pourrait fournir des éléments clés de compréhension de cette période.

Une première approche a considéré le réseau sans la composante temporelle. Pour cela, nous avons proposé dans [1] un modèle stochastique à groupes latents qui considère la partition en royaumes et provinces du réseau comme une décomposition en sous-graphes. Chaque sous-graphe a son propre vecteur de mélange et voit ses sommets associés aux clusters. Les sommets sont alors connectés avec une probabilité dépendant des sous-graphes seulement, alors que les types des arêtes sont supposés être échantillonnés à partir des groupes latents. L'algorithme *variational Bayes EM* est proposé pour l'inférence ainsi qu'un critère de sélection de modèles pour l'estimation du nombre de cluster.

Afin de prendre en compte l'aspect temporel (qui apparait clairement sur la figure 1), nous proposons d'étendre le modèle RSM en modélisant l'évolution temporelle des noeuds et de leur connection grâce à un *state space model*. Ce modèle permet de lier les probabilités de connection des noeuds à l'instant t à des paramètres d'état dans un espace latent au travers d'une transformation logistique.

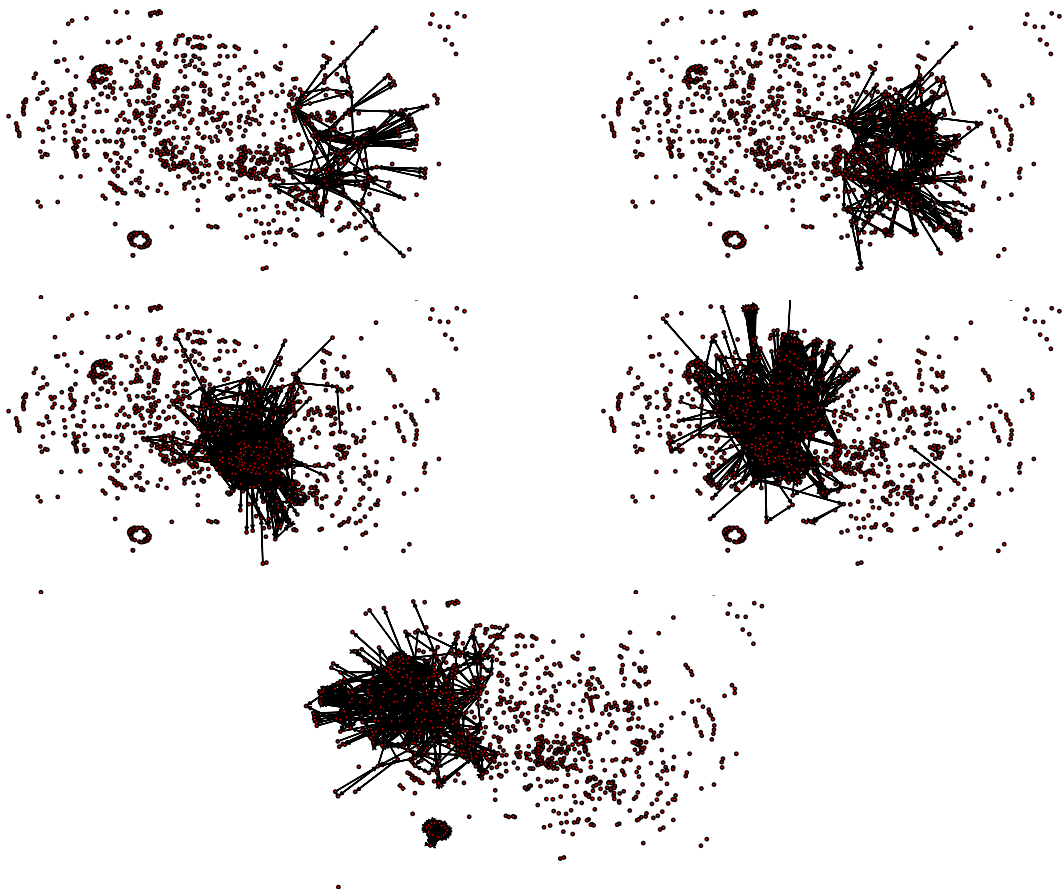


Figure 1: Dynamique du réseau des évêques par tranche de 30 années entre 500 et 650 après J.C.

Bibliographie

[1] Y. Jernite, P. Latouche, C. Bouveyron, P. Rivera, L. Jegou & S. Lamassé, *The random subgraph model for the analysis of an ecclesiastical network in merovingian Gaul*, The Annals of Applied Statistics, in press, 2014.