



**ined**  
INSTITUT  
NATIONAL  
D'ÉTUDES  
DÉMOGRA-  
PHIQUES

Séminaire du service méthodes statistiques de l'Ined  
en partenariat avec le GDR  
"Analyse de réseaux en sciences humaines et sociales"



## LES RENCONTRES DE STATISTIQUE APPLIQUÉE

# Analyses statistiques de réseaux

Vendredi 15 septembre 2017

9h30 - 17h (Salle Alfred Sauvy)

L'analyse de réseaux s'impose aujourd'hui dans de nombreux domaines pour analyser, visualiser et modéliser les données. Ces dernières sont devenues de plus en plus faciles à obtenir et massives, elles restent cependant souvent peu structurées et hétérogènes. Dans plusieurs disciplines, ces techniques permettent d'explorer un corpus dans toutes ses dimensions. Les exposés de cette session seront centrés sur des éléments de méthodes, ou présenteront des outils (logiciel Puck pour l'analyse des réseaux de parenté, site linkage pour des réseaux textuels, visualisations multiples de mesures de la production scientifiques).

En savoir plus sur <https://statapp.site.ined.fr/>

### Résumés

**9h30 : Arnaud Bringé (Ined, SMS) Laurent Beauguitte (CNRS, UMR IDEES) et Karl Van Meter (ENS)**  
**Accueil de la journée**

**10h : Thomas Louail (CNRS, UMR Géographie-cités)**  
**Analyse des réseaux de la mobilité urbaine**

L'expression "réseaux de la mobilité urbaine" évoque immédiatement les réseaux de transport, les infrastructures : les rues, les lignes de métro, bus, tram, dont les propriétés sont étudiées depuis longtemps déjà avec les métriques et modèles de la théorie des graphes et de l'analyse statistique des grands réseaux. Depuis une dizaine d'années on dispose également de plus en plus de données au sujet des déplacements "sur" ces réseaux : flux et trajectoires individuelles, lesquelles se prêtent aussi à une représentation par des réseaux. L'expression évoque enfin un troisième type de réseaux, ceux de la co-présence spatiale, leurs propriétés dynamiques, lesquelles découlent des propriétés de la mobilité individuelle et des relations sociales.

Une liste non-exhaustive de types de questions qui ont été étudiées au sujet de ces trois types de réseaux et quelques résultats statistiques de la littérature de la décennie passée seront proposés ici.

**10h35 : Julien Brailly (Ecole Nationale Supérieure Agronomique de Toulouse (INP-ENSAT))**  
**Etudier l'influence sociale dans un réseau par les modèles ALAAM : la diffusion d'une norme de punitivité au sein d'une communauté agropastorale sénégalaise**  
(A venir)

**11h25 : Fabien Eloire (Univ. Lille 1)**

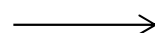
**La triade comme sous-structure élémentaire pour l'analyse des processus sociaux**

La statistique des réseaux sociaux, contrairement à la statistique classique, fait face à l'épineux problème de la dépendance des observations. Pour le résoudre, elle propose un nouveau type de modélisation statistique, les *Exponential Random Graph Models* (ERGM). À la base de ces modèles, se trouve la notion élémentaire de triade, discutée de longue date dans la littérature sociologique. La communication met en avant l'intérêt de l'approche triadique pour l'analyse des processus sociaux.

**12h : Camille Roth (Sciences Po (médiab), Centre Marc Bloch e.V.)**

**Calibration et validation des modèles dynamiques de réseaux sociaux**

La formalisation et l'estimation empirique des dynamiques interactionnelles joue un rôle crucial dans la modélisation sociale: au niveau micro, des méthodes statistiques sophistiquées ont été introduites pour appréhender les préférences de création de liens; au niveau macro, les modèles de réseaux sociaux formulent des hypothèses empiriques ou normatives afin de lier formation de liens et motifs structurels globaux. Autrement dit, les interactions sont soit la cible soit le fondement de ces modèles. Cet exposé proposera d'abord un tour d'horizon de la littérature consacrée à ce type de modèles, puis introduira une technique de régression symbolique permettant d'appréhender conjointement estimation et modélisation des dynamiques d'interaction globales et locales.



**14h15 : Marion Maisonobe (féd. INCREASE (FR CNRS 3707), Poitiers-UMR LISST, Univ. Toulouse)**

**La représentation et l'analyse des réseaux scientifiques mondiaux**

*(A venir)*

**14h50 : Pierre Latouche (SAMM - Univ. Paris 1)**

**Linkage : analyse conjointe de réseaux et de textes**

La plupart des réseaux sociaux décrivent des échanges entre des individus, échanges se faisant par l'intermédiaire de textes. C'est en particulier le cas des réseaux de type Facebook ou Twitter, mais également des réseaux d'emails. Or, de manière tout à fait surprenante, les outils d'analyse de réseaux, qui ont pourtant été souvent motivés par des questions pratiques en sciences sociales, modélisent pour la plupart ces réseaux de manière binaire. Autrement dit, l'information de texte échangé disparaît et seules les présences ou absences de connexions entre individus sont retenues comme source de données. Nous proposons donc le stochastic topic block model ainsi qu'une procédure d'inférence associée permettant l'analyse conjointe d'un réseau dynamique et d'un ensemble de textes portés par les connexions du réseau. Nous mettons en avant l'utilisation d'une variable pivot autorisant les analyses alternées du réseau et du corpus associé. Un critère de sélection de modèles est dérivé afin d'estimer le nombre de thèmes de discussions et le nombre de clusters d'individus. Nous illustrons nos résultats sur des données simulées ainsi que sur la base de données d'emails du scandale Enron. Le cœur de ce travail est implémenté dans la plateforme linkage.fr.

**15h40 : Christophe Prieur (Télécom ParisTech)**

**Fouille de données pour étudier les sociabilités**

Avec l'évolution des outils de communication, les manières de tisser et entretenir des liens se transforment, en même temps que les méthodes qu'il est possible de mettre en œuvre pour étudier ces liens. Les sciences sociales dites computationnelles s'appuient sur des très grandes quantités de données pour étudier des phénomènes sociaux. Si on les applique en gardant à l'esprit la nécessité de suivre de près les situations individuelles, ces méthodes peuvent permettre d'enrichir très significativement l'analyse d'une grande variété de comportements. J'illustrerai ce genre de démarche à partir d'une enquête réalisée sur les usages de Facebook.

**16h15 : Pascal Cristofoli (EHESS,LaDeHiS)**

**Mesures de réseaux et parenté avec le logiciel Puck**

L'exposé commencera par une présentation de la fonctionnalité principale du logiciel Puck (Program for the Use and Computation of Kinship data), le recensement des circuits matrimoniaux dans un corpus généalogique. Il reviendra sur les particularités des réseaux de parenté, les concepts mobilisés, les modalités de paramétrage de l'opération de recensement et les différents résultats proposés. Il abordera les questions cruciales corollaires à l'interprétation de ces résultats: la nécessaire connaissance des biais d'un corpus de parenté et celle de la représentativité des résultats obtenus et les perspectives comparatives. Enfin, il présentera les évolutions récentes du logiciel et les opportunités ouvertes par l'extension de son modèle de données qui permet désormais de prendre en compte toutes sorte de données relationnelles.

**16h50 : Discussion et conclusion**

Ce séminaire est ouvert à tous sans frais de participation mais sur inscription à :

<http://www.ined.fr/fr/actualites/rencontres-scientifiques/seminaires-colloques-ined/rsa-15-09-17/>

INED : 133, boulevard Davout, Paris 20e • Standard: 01 56 06 20 00

Métro : L9 (Porte de Montreuil) ou L3 (Porte de Bagnolet) Bus 57 ou Tram T3b (Marie de Miribel)