

Résultats de théorie des graphes sous-jacents à l'analyse de réseaux réels interdisciplinaires.

Romain BOULET

Université Lyon 3 / IAE

Equipe Magellan

www.romainboulet.fr

Atelier PROSPECTOM

Grenoble

Jeudi 29 novembre 2012

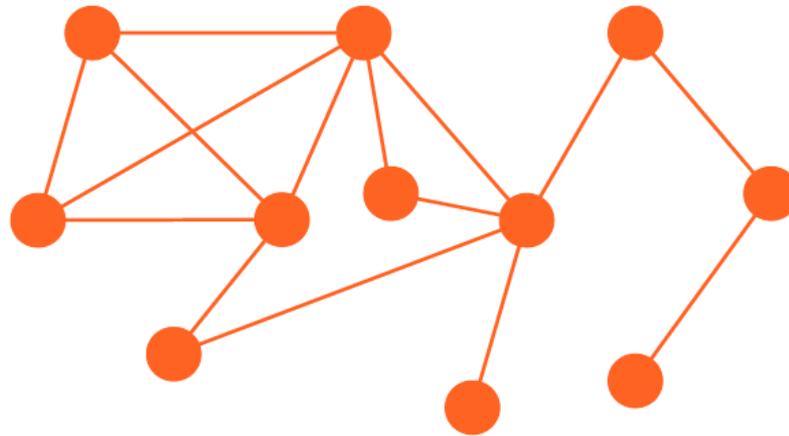
Plan

- Introduction : le graphe, les réseaux et l'interdisciplinarité
- Graphes, réseaux et interactions interdisciplinaires
 - Réseaux sociaux entre paysans au Moyen-âge
 - Reconstruction de cartes
 - Réseaux de textes juridiques
 - Thésaurus
- Perspectives : réseaux en sciences biologiques

Introduction : Le graphe, les réseaux et l'interdisciplinarité

Définition

- ▶ Un graphe est :
 - Un ensemble V de sommets
 - Un ensemble E (inclus dans $V \times V$) d'arêtes



Propice à l'interdisciplinarité

Sociologie

- Réseaux sociaux

Biologie

- Réseaux de protéines

Chimie

- Modélisation de molécules

Sciences informatiques

- Réseaux de serveurs, *cloud*

Sciences économiques

- Réseaux entre entreprises du CAC40

Linguistique

- Ontologies

etc.

Graphes, réseaux et interactions interdisciplinaires : quelques exemples

I – Réseau social entre paysans du moyen-âge

~~1277~~ 1277. 10. avril
N° 68 in bruta

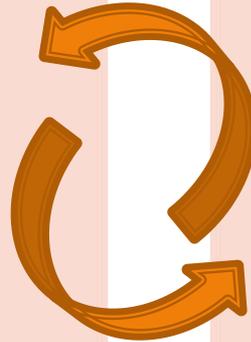
Scute par Bernard de balandran fe de ar de
balandran le B. et ar. ses fils.
a Jean et a G. de montignac freres.
D'une piece de terre au sud de la paroisse de ysom
tenant d'une part terre de J. et G. de montignac
freres d'autre part terre de arnal de vivier, d'autre part
terre de B. de la garrigue
pour le prix de trente sols.
La piece de terre tenue en fief du chapitre de laon
guillaume de fabrica use de faitllesse

<http://graphcomp.univ-tlse2.fr/graph>

Enrichissement interdisciplinaire

Théorie des graphes

- Détection de communautés particulières via le Laplacien
- Mise en place d'une visualisation simplifiée du réseau



Histoire médiévale

- Architecture des liens sociaux entre paysans au Moyen-âge
- Visualiser et comparer les réseaux sociaux paysans aux Moyen-âge

Communautés parfaites

Définition

Une *communauté parfaite* est un sous graphe complet dont les sommets ont le même voisinage fermé.

- → Notion de module
- → Notion d'équivalence structurale

Théorème

Un ensemble non stable S de sommets d'un graphe G d'ordre n est une communauté parfaite si et seulement s'il existe une valeur propre non nulle du Laplacien de G de multiplicité au moins $k - 1$ ($k > 1$) telle qu'il existe $k - 1$ vecteurs propres ayant les mêmes $n - k$ coordonnées nulles.

- → Détection des communautés parfaites via les éléments spectraux.

Communautés parfaites

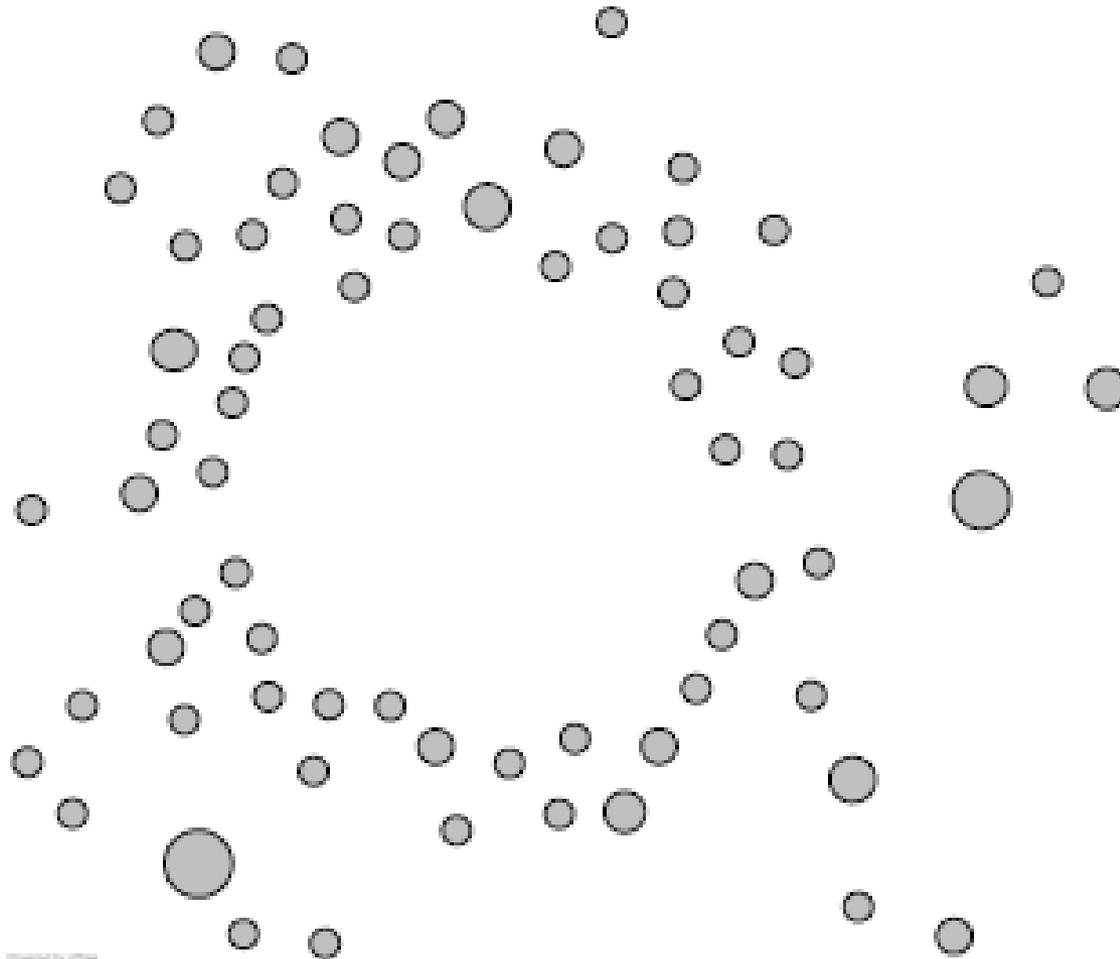


Illustration by J. J. P.

Liens entre deux communautés

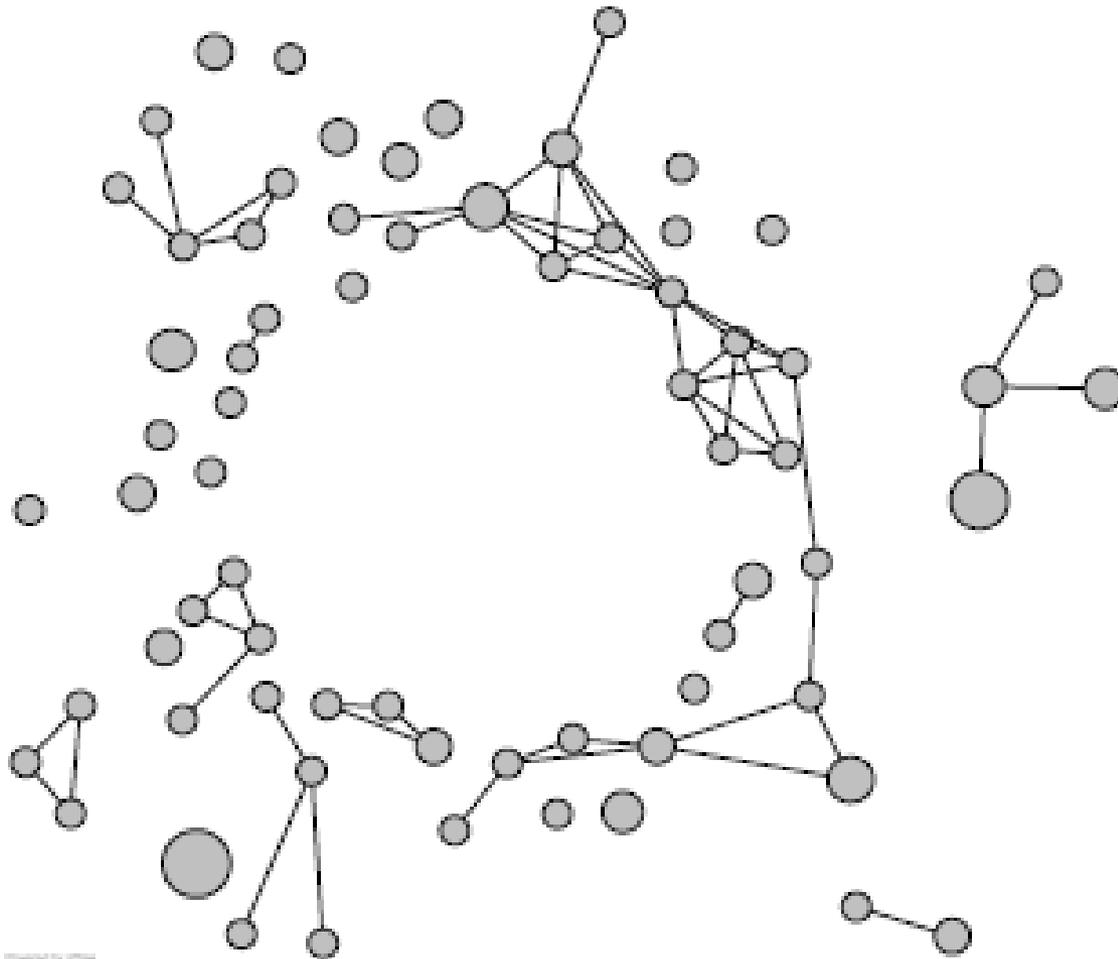


Illustration by J. J. J.

Présence d'un club-huppé

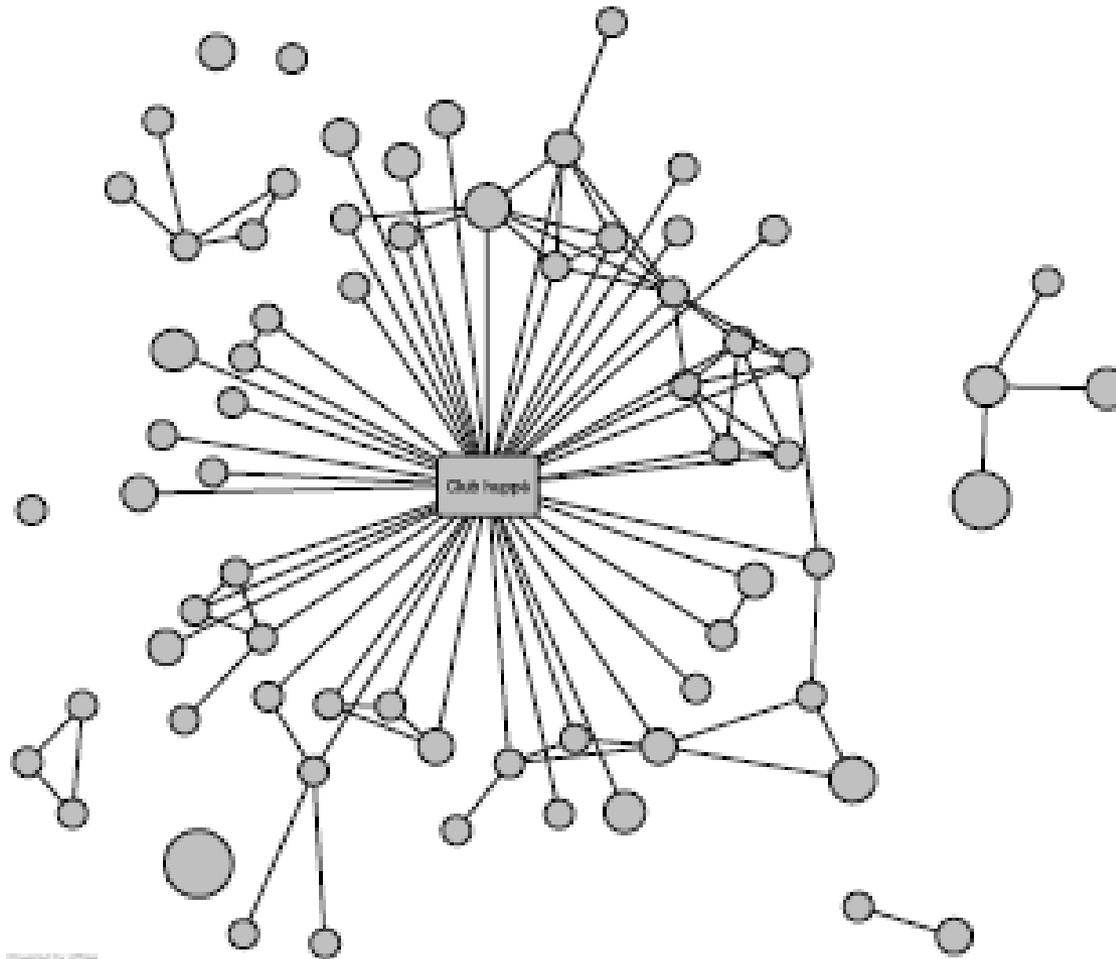


Illustration by J. J. J.

Présence de sommets centraux

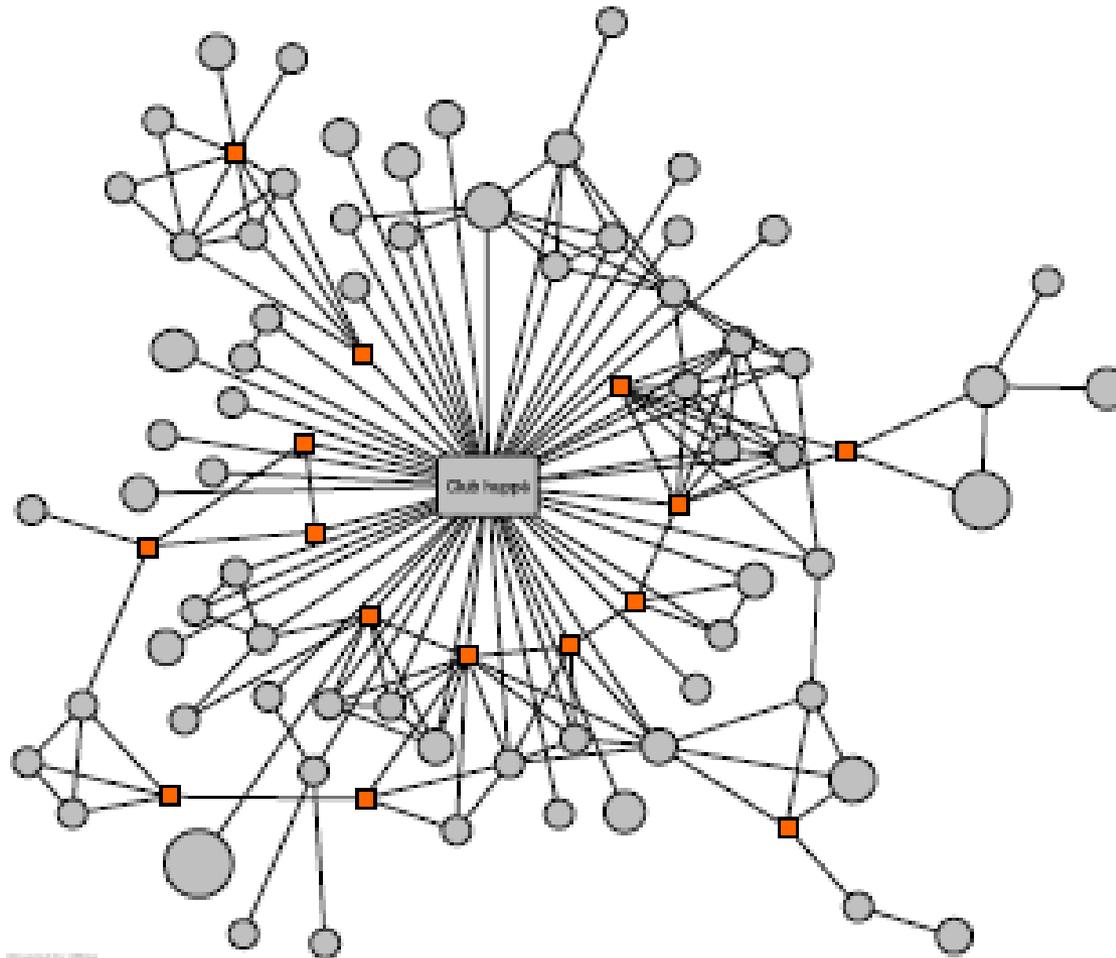
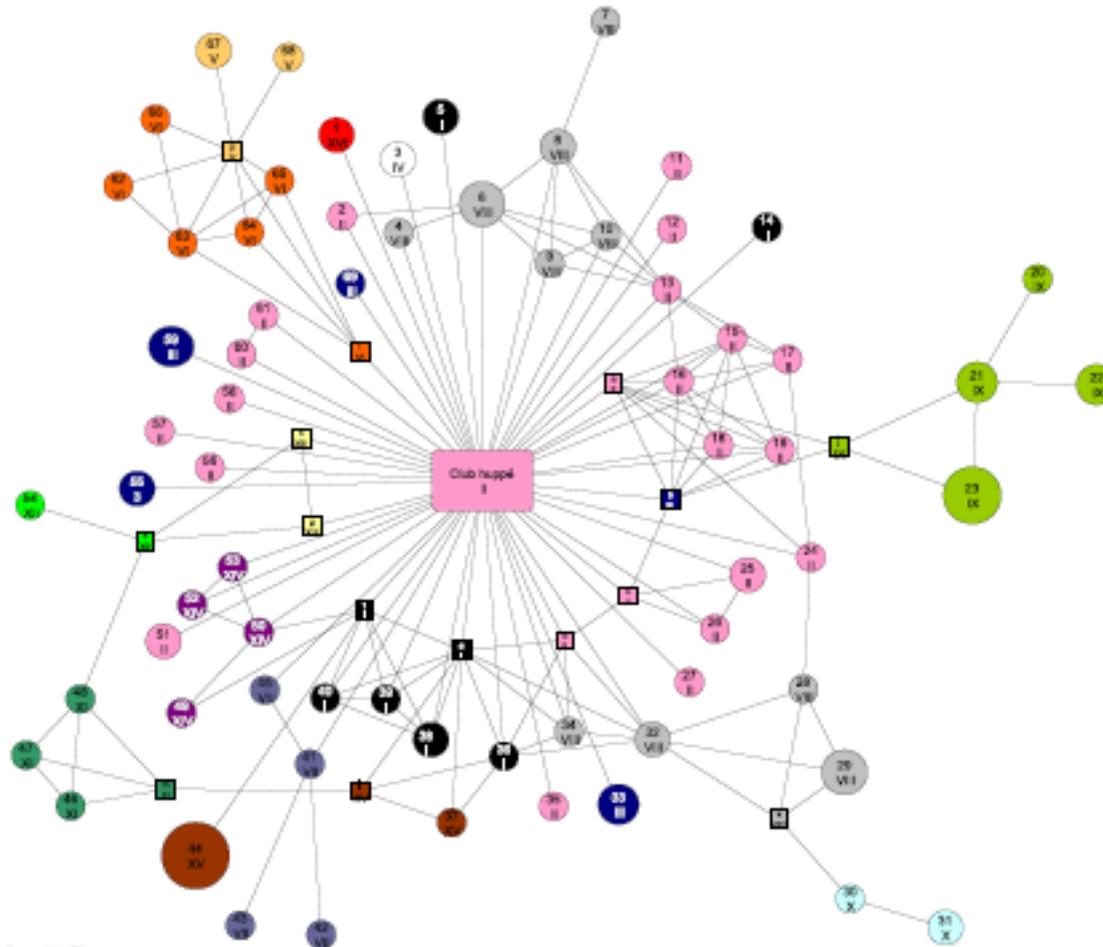
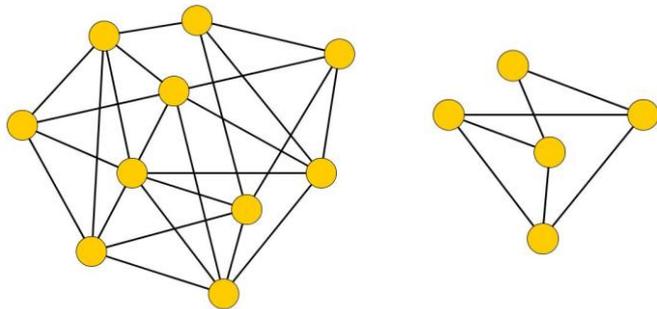


Illustration by J. J. J.

Regroupement via partitionnement spectral



Partitionnement spectral

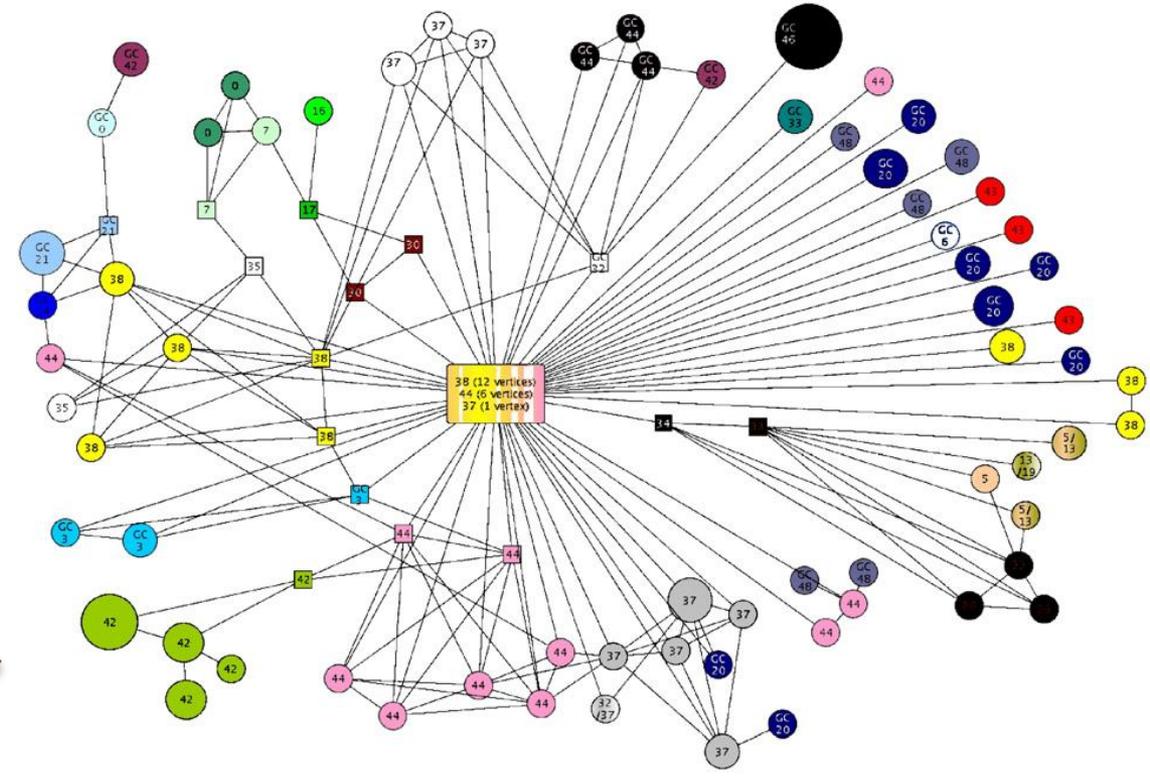


1	0
1	0
1	0
1	0
⋮	⋮
0	1
0	1
0	1

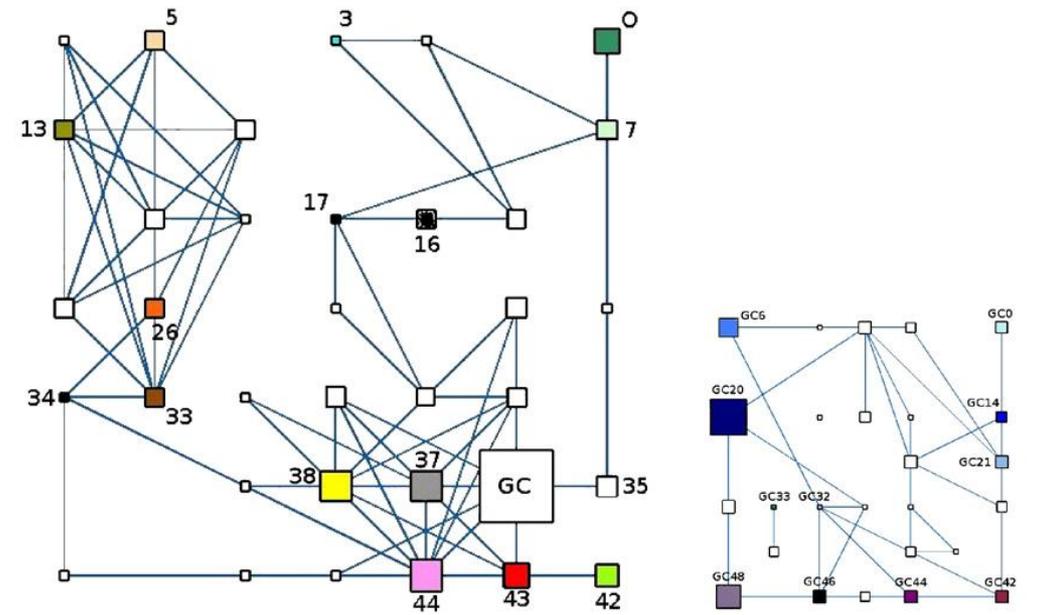
- ▶ $L=Q\Lambda Q'$
- ▶ Les k premiers éléments propres plongent les sommets dans R^k
- ▶ On effectue une classification dans R^k du type k -means.

Cartes de Kohonen

Visualisation de communautés obtenues via des propriétés spectrales de la matrice laplacienne



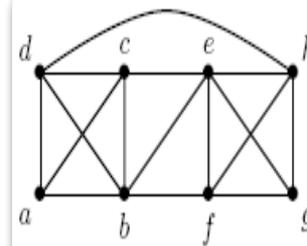
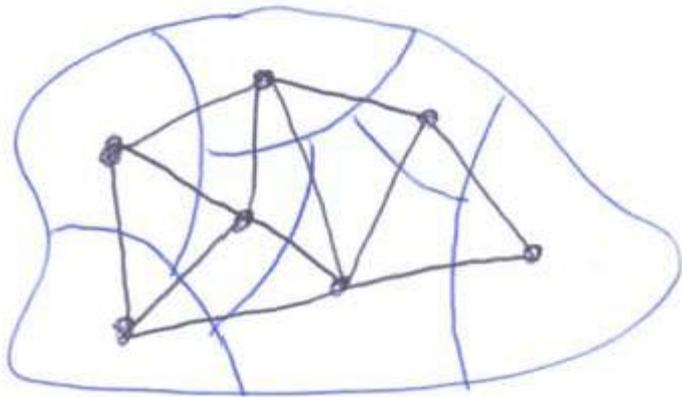
Utilisation d'un algorithme de cartes auto-organisatrices sur ces données.



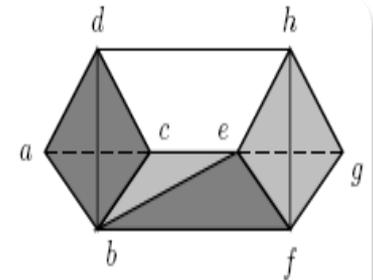
Bibliographie

- ▶ Boulet R., Jouve B., Rossi F., Villa N., **Batch kernel SOM and related Laplacian methods for social network analysis**, Neurocomputing 71(7–9), 1257–1273
- ▶ Boulet R., Jouve B., **Partitionnement d'un réseau de sociabilité à fort coefficient de clustering**, Revue des Nouvelles Technologies de l'Information (9), 569–574.
- ▶ Villa N., Boulet R., **Clustering a medieval social network by SOM using a kernel based distance measure**, In Proceedings of ESANN 2007, Bruges, Belgique, 31–36.
- ▶ Boulet R., et al , **Sur l'analyse de réseaux de sociabilité dans la société paysanne médiévale**. Actes des journées MASHS 2007, Brest, France, 10–11 Mai 2007.
- ▶ Boulet R., **Comparaison de graphes, applications à l'étude d'un réseau de sociabilité paysan au Moyen Âge** (200 pp). Thèse de doctorat

II – Reconstruction de cartes



H

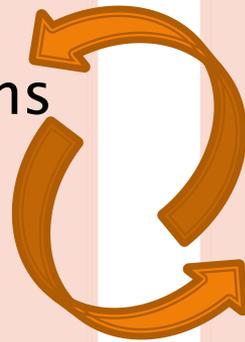


$\Delta_g(H)$

Enrichissement interdisciplinaire

Théorie des graphes

- Détecter les « trous » dans un graphe triangulé
- Etudier les déformations de graphes
- Lien entre graphe et topologie



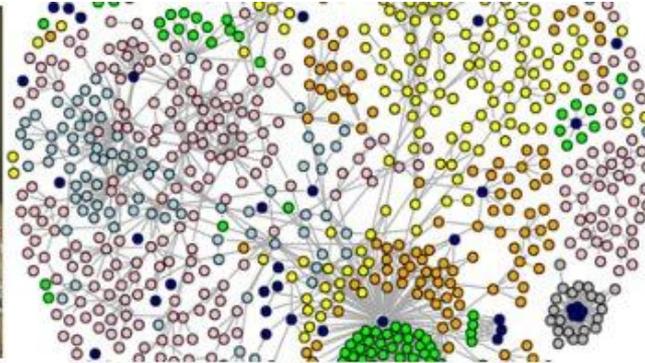
Sciences historiques

- Reconstruire des cartes à partir d'écrits
- Détecter les parcelles non décrites
- Etudier l'évolution de ces cartes

Bibliographie

- ▶ Boulet R, Fieux E, Jouve B
s-homotopy for finite graph
Topological & Geometrical Graph Theory TGGT2008,
Electronic Notes in Discrete Mathematics 31 (2008), 123–127.
- ▶ Boulet R., Fieux E., Jouve B.
Simplicial simple-homotopy of flag complexes in terms of graphs.
European Journal of Combinatorics, 31 (2010), pp. 161–176

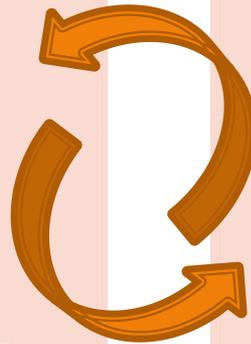
III – Réseaux de textes juridiques



Enrichissement interdisciplinaire

Théorie des graphes

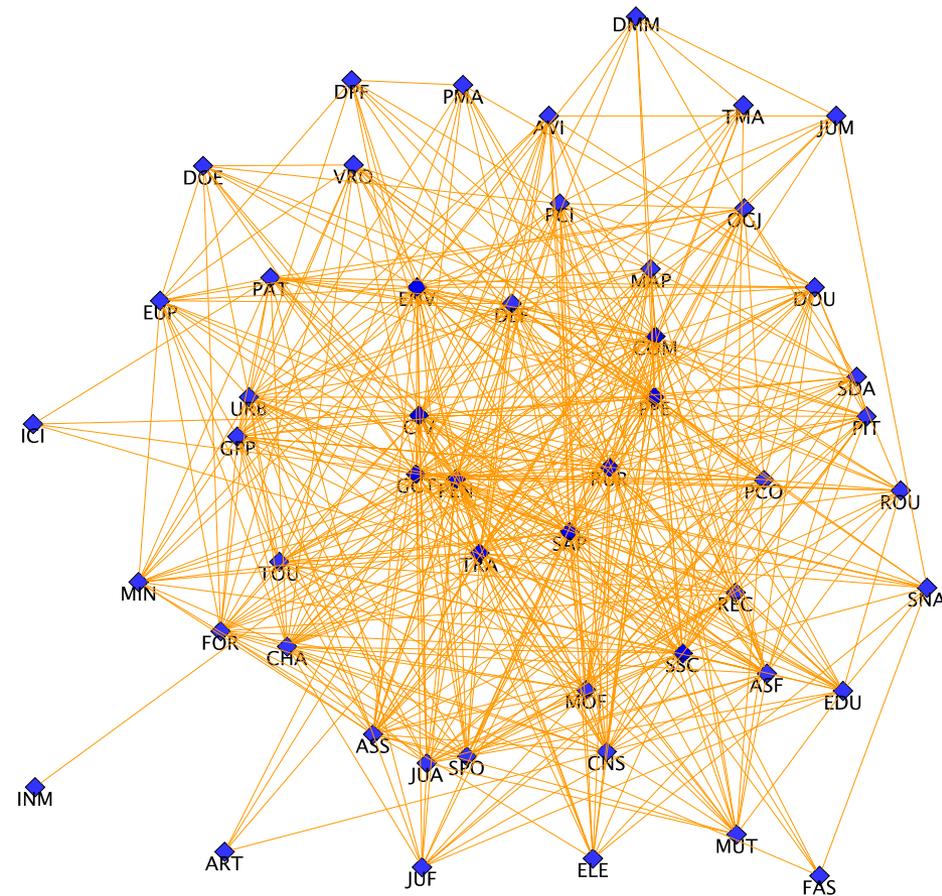
- Analyse d'un réseau dense
- Représentation alternative de ce réseau



Sciences juridiques

- Etablir une cartographie des grands domaines juridiques

Réseau des codes législatifs français



LCA

► Un réseau dense

► Un « faux semblant » de petit-monde

↳ Un monde concentré

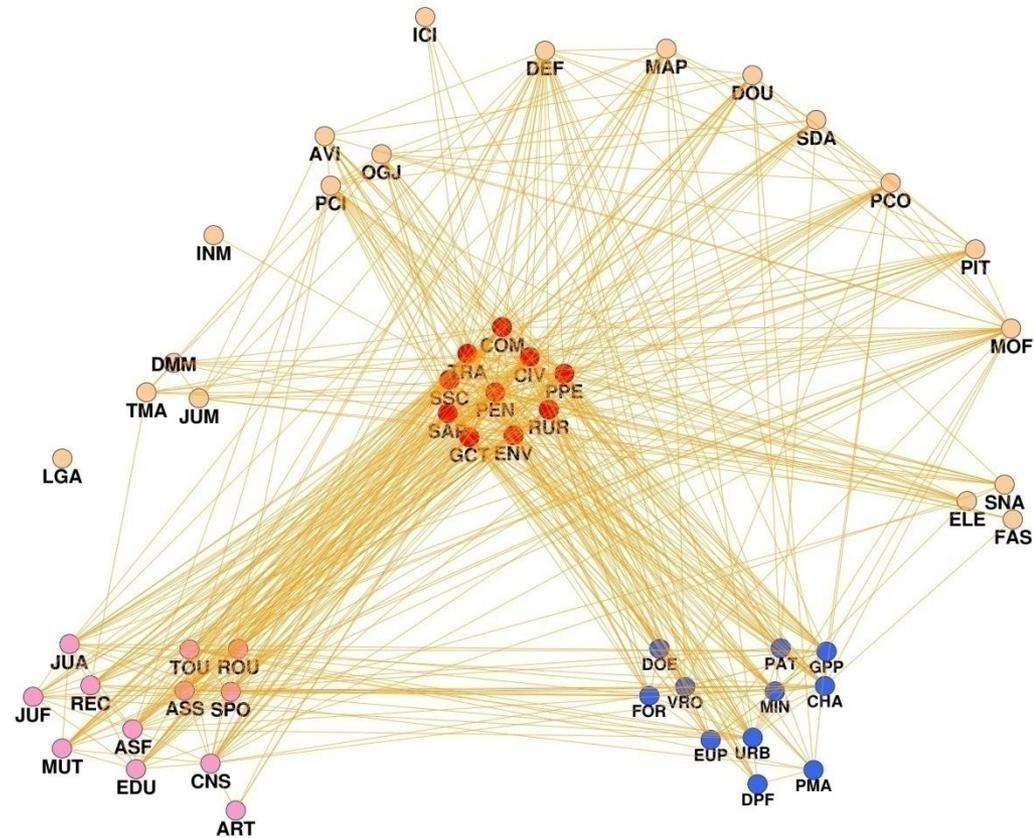
- forte concentration d'arêtes
- centralités concentrées dans 10 sommets

Méthode d'analyse

- On construit le réseau + analyse structurale
- On détecte et enlève les *club-huppé* (dédensification).
- On partitionne le graphe résultant via trois algorithmes de partitionnement.
- Nous extrayons les communautés stables.
- Nous ajoutons le *club-huppé* dans la représentation finale.

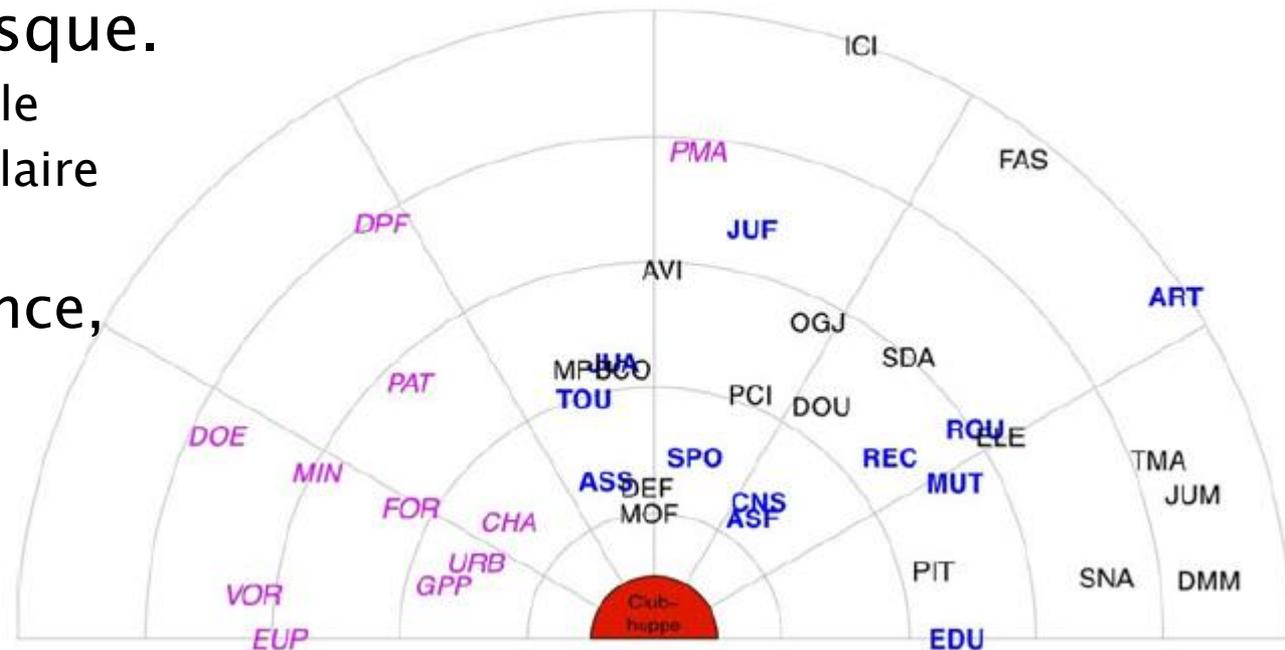
Une première cartographie

- ▶ Groupe central et influent de codes
 - Pénal, Civil, Environnement, ...
- ▶ Groupe de codes à caractère social
 - Education, Consommation, Sport
- ▶ Groupe de code relevant du foncier
 - Expropriation, Minier, Patrimoine



Représentation en hémicycle

- ▶ Représentation alternative
 - Sans les arêtes
 - Orienté « juriste »
- ▶ On représente les sommets sur un demi-disque.
 - Composante radiale
 - Composante angulaire
- ▶ Notions de distance, de réduction de dimension, ...



Fondations mathématiques de « l'hémicycle »

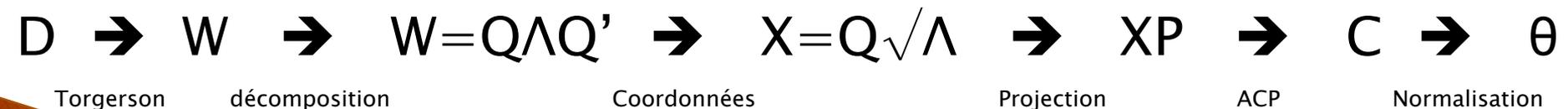
- ▶ On munit l'ensemble des sommets du graphe de la dissimilarité (euclidienne) de Dice :

$$\delta^2(v, w) = \frac{|\Gamma_v \Delta \Gamma_w|}{|\Gamma_v| + |\Gamma_w|}$$

- ▶ On a la distance moyenne au club huppé d'un sommet :

$$\bar{\delta}(j) = \frac{1}{|\text{Rich} - \text{Club}|} \sum_{k \in \text{Rich} - \text{Club}} \delta(k, j)$$

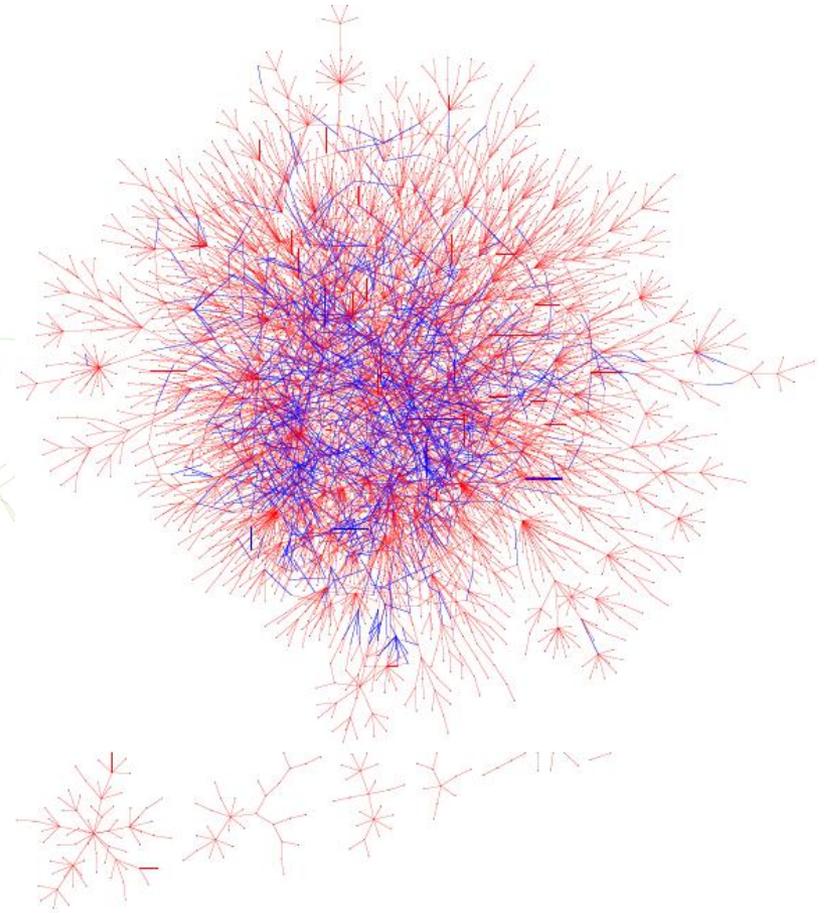
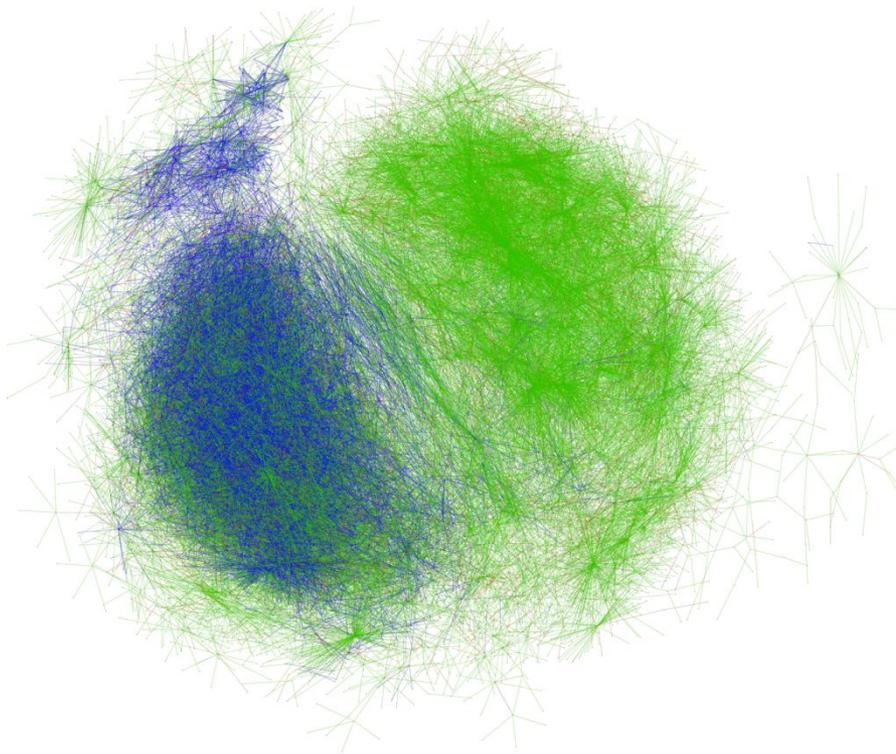
- ▶ On calcule la composante angulaire :



Bibliographie

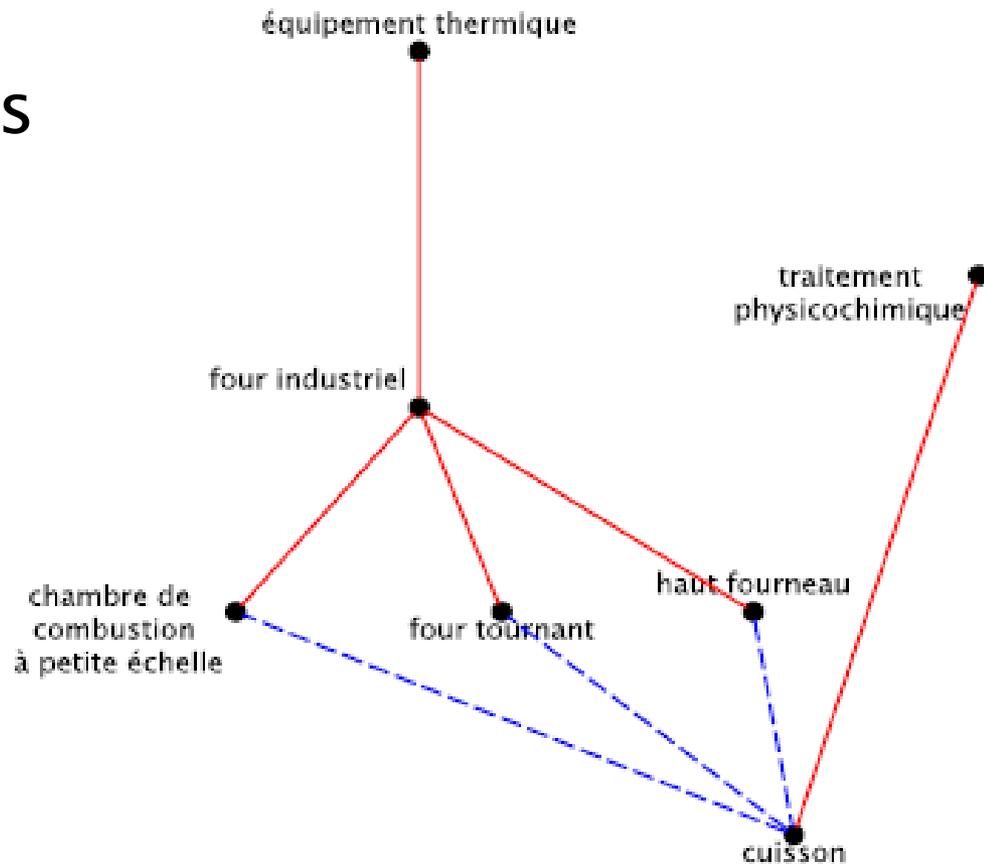
- ▶ Boulet R., Mazzega P. and D. Bourcier **A Network Approach to the French System of Legal codes – Part I: Analysis of a Dense Network**. *Artificial Intelligence and Law*, 19, 333–355 (2011).
- ▶ Bourcier D., Mazzega P., Boulet R., **Visualiser la Complexité du Droit**, Les technologies de l'information au service des droits : opportunité, défis, limites sous la coordination de Daniel Le Métayer, Bruylant, Bruxelles, (2010)
- ▶ Boulet R., Mazzega P., Bourcier D., **Réseaux normatifs relatifs à l'environnement : structures et changement d'échelles**. Atelier Complexité et Politiques Publiques, Paris, septembre 2010.
- ▶ Boulet R., Mazzega P., Bourcier D. **Network Analysis of the French Environmental Code**, Artificial Intelligence approaches to the complexity of legal systems (AICOL 2009, Rotterdam)
- ▶ Boulet R., Mazzega P., Jouve B., **Environmental Social and Normative Networks in the MAELIA Platform**, Workshop on Legal and Negotiation Decision Support Systems (LDSS 2009 Barcelone), M. Poblet, U. Schild, J. Zeleznikow eds., CEUR, p.83–93.
- ▶ Mazzega P., Bourcier D., Boulet R., **The Network of French Legal Codes**, Twelfth International Conference on Artificial Intelligence and Law (ICAIL 2009, Barcelone).
- ▶ Mazzega P., Bourcier D., Boulet R., **Code Communities in the French Legal System**, Symposium on Courts and Mediations – New Paths for Justice. 18–19 June, 2009.

IV- Thésaurus



Des réseaux multiplexes

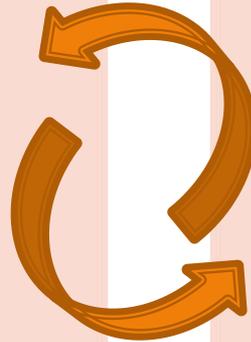
- ▶ Réseau multiplexe :
réseau avec différents types de liens
- ▶ Pour un thésaurus :
 - Liens hiérarchiques (**narrower/broader**)
 - Liens d'association (**related to**)



Enrichissement interdisciplinaire

Théorie des graphes

- Analyse de réseaux multiplexes



Recherche d'Information

- Comprendre l'organisation des thésaurus
- Amélioration des moteurs de recherche

Analyse de thésaurus

Analyse structurelle

- Indices de petit-monde
- Indices de centralité
- Distribution des degrés

7 thésaurus analysés

Détection de communautés

- Evaluation de plusieurs algorithmes de partitionnement
- Visualisation de communautés

Problématique des graphes multiplexes

Navigation entre communautés

- Intégration dans un moteur de R.I.

Des indices de structures sur graphes multiplexes

- ▶ Mise en place d'une batterie d'indices pour mesurer :
 - Gain de connexité
 - Évolution de connectivité globale
 - Connectivité locale (triangles hétérogènes)
 - Répartition dans les communautés

- ▶ Afin de :
 - Obtenir des modèles, une typologie
 - Guider l'édition des thésaurus

Exemple d'indice : connectivité locale

- ▶ Graphe simple : connectivité locale mesurée par la probabilité que deux sommets soient liés sachant qu'ils ont un voisin commun :

$$6 \times (\text{nb de triangles}) / \text{nb de triplets connexes}$$

↳ taux de transitivité

- ▶ graphe multiplexe : apport des liens **related to** dans ce taux de transitivité

Exemple d'indice : connectivité locale

- ▶ Calcul d'un indice de transitivité (via la spectre de matrice d'adjacence)

$$\text{tr}((A_n + \zeta A_r)^3) = \text{tr}(A_n^3) + \zeta \text{tr}(A_n^2 A_r + A_r A_n^2 + A_n A_r A_n) \\ + \zeta^2 \text{tr}(A_r^2 A_n + A_n A_r^2 + A_r A_n A_r) + \zeta^3 \text{tr}(A_r^3)$$

$$\sum \lambda^3 = 6 \times \triangle + \zeta \times 6 \times \triangle + \zeta^2 \times 6 \times \triangle + \zeta^3 \times 6 \times \triangle$$

où $\zeta = e^{\frac{2i\pi}{5}}$

Nombre de \triangle :

GEMET	Modèle aléatoire
38	2.27

Bibliographie

- ▶ Romain BOULET, **Introduction d'indices structuraux pour l'analyse de réseaux multiplexes**. Conférence MARAMI 2011, Grenoble (19–21 octobre 2011).
- ▶ Mazzega P., Boulet R. and Libourel Th. **Graphs for Ontology, Law and Policy** in *"New Frontiers in Graph Theory"*, ISBN 979-953-307-303-2 (2012).
- ▶ Sayah, H. **Composant Sémantique pour l'amélioration de la recherche de données environnementales** (stage de Master 2 Informatique, Université Montpellier 2)

Perspectives : réseaux en sciences biologiques

Théorie des graphes et bioinformatique

Arbres phylogénétiques

Classification des espèces / ontologie

Voies métaboliques

Réseau d'interactions entre protéines

Réseaux de régulations / interactions de gènes

Protéines (chaînes d'acides aminés)

Enrichissement interdisciplinaire

Théorie des graphes

- Analyse de *grands graphes*
- Développement de méthodes spectrales



Bio-informatique

- Obtention de « communautés »
- Recherche de motif à l'intérieur d'un réseau

Analyse de la complexité (de réseaux)

Dialogue interdisciplinaire

Métrieologie statique
Structure générale, communautés

Multiplexité
Interactions des différents types de liens (diverses relations)

Dynamique
Vision temporelle (dynamique des réseaux)

Emergence, Méta-réseau
Compréhension de la complexité, effets d'échelle (méta-réseau)

**Merci pour votre
attention !**