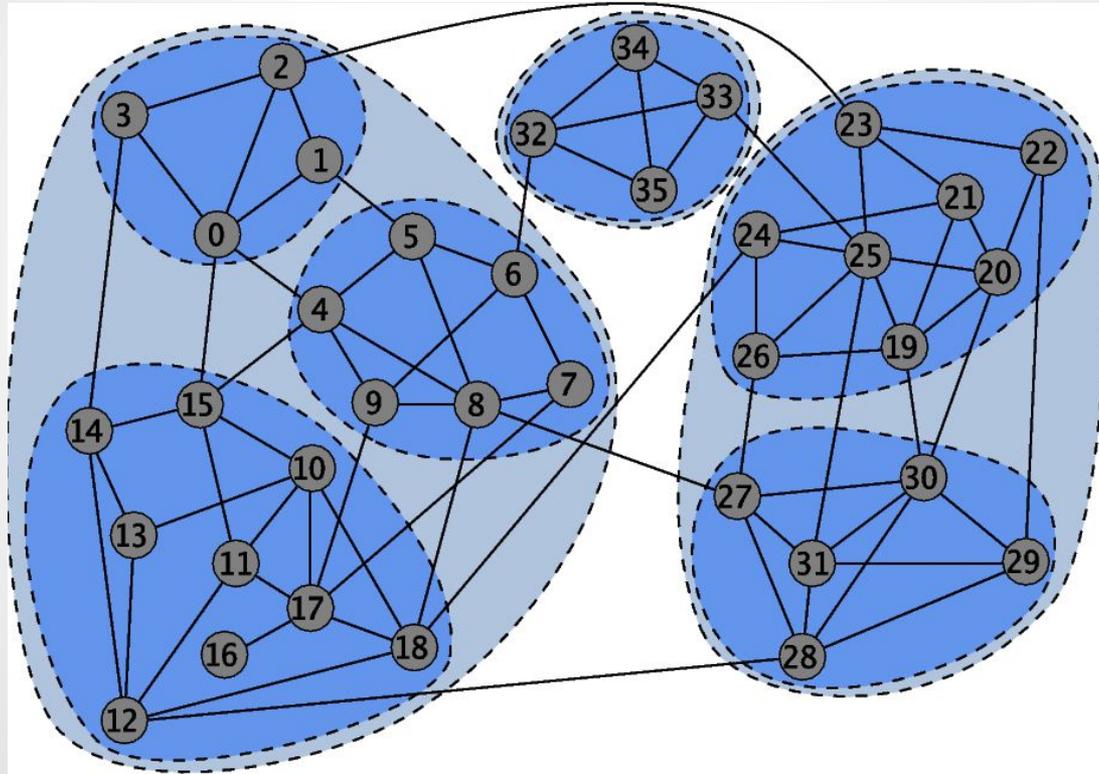


Détection de communautés dans les flots de liens par optimisation de la modularité.

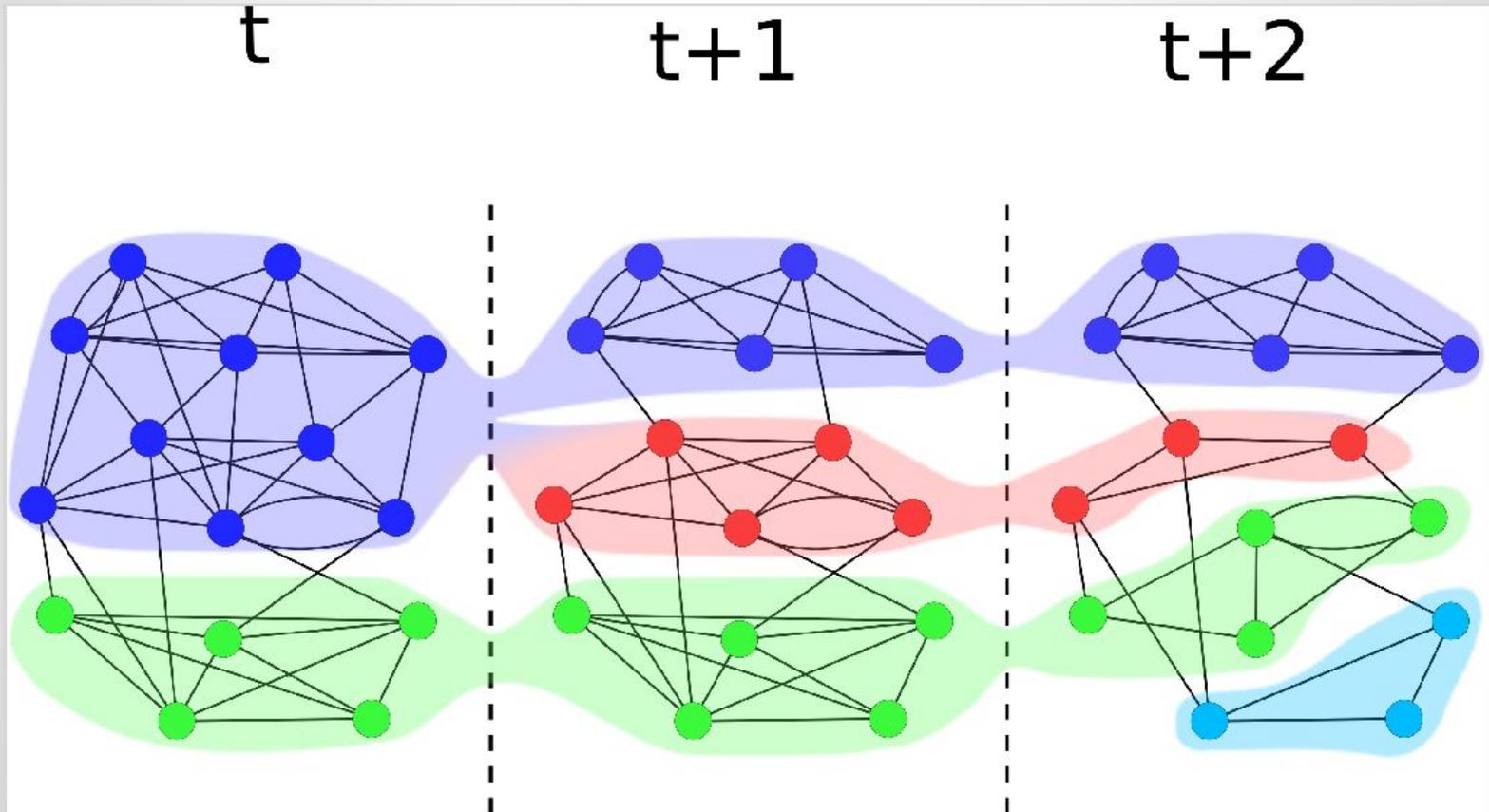
Emmanuel Orsini

LIP6

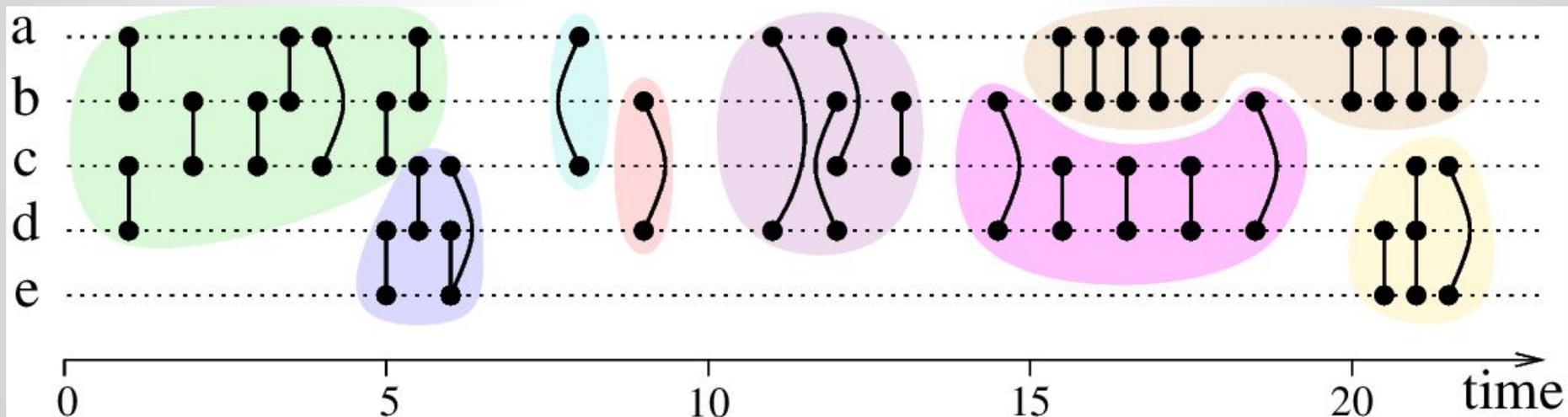
Problématique



Problématique



Problématique



Modularité d'un graphe statique

e_{ij} = proportion d'arêtes entre les communautés C_i et C_j

a_i = proportion de degré dans C_i

$$Q = \sum e_{ii} - a_i^2$$

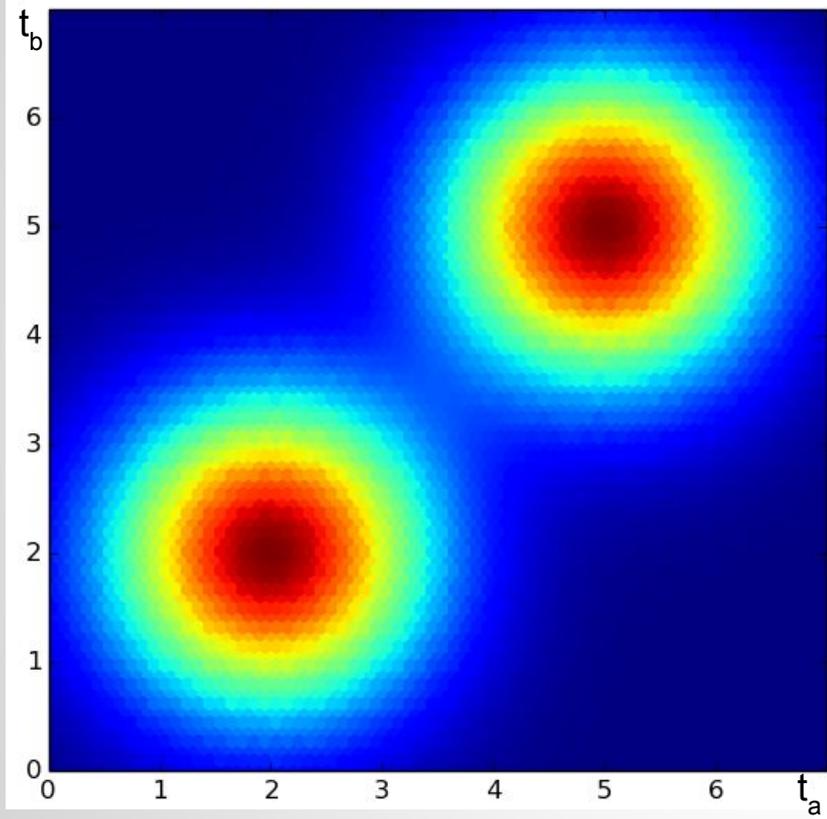
Comment donner un sens à la modularité dans les flots de liens?

- Possibilité d'agréger les noeuds
- Les communautés doivent être connexes sinon elles seront séparées

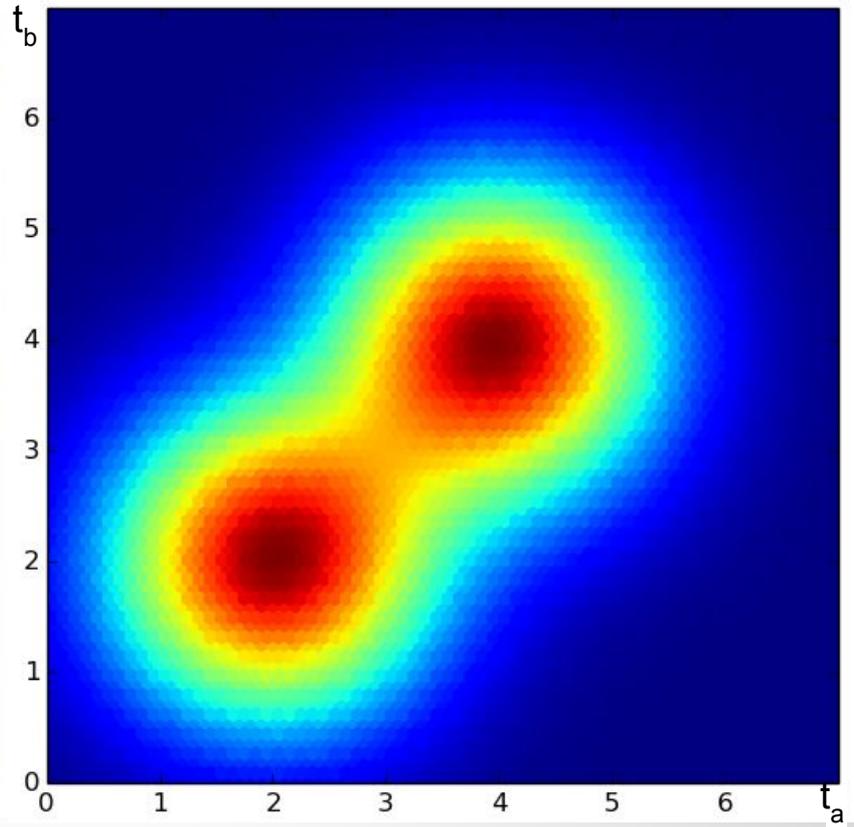
Création d'un nouvel objet

- Considérer comme indépendant le noeud à chaque instant: on obtient les noeuds datés
- Construire une fonction d'adjacence entre les noeuds datés qui conservent les propriétés de communautés

La fonction d'adjacence



Liens à $t = 2$ et $t = 5$



Liens à $t = 2$ et $t = 4$

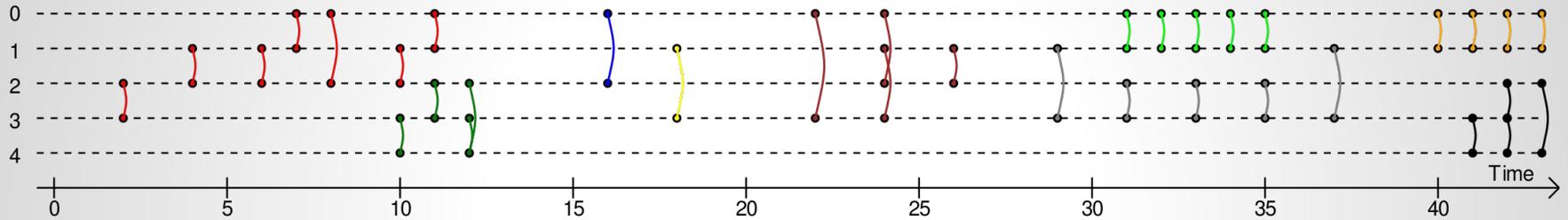
Mise en pratique

Le graphe obtenu a trop de noeuds, il faut opérer une première étape d'agrégation pour alléger le graphe.

→ Par exemple borner le nombre de noeuds par le nombre d'interactions.

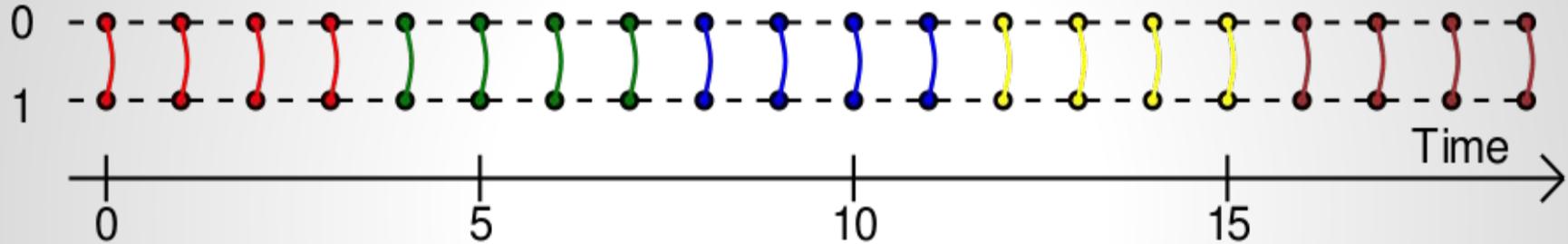
→ Complexité linéaire par rapport aux nombre de liens.

Résultats obtenus



Sur de vrais jeux de données:
400 000 mails en 30 minutes

Limitation



En réalité, on a construit un graphe qui n'a pas tout à fait une structure communautaire!

Apport à l'état de l'art

- Une nouvelle modélisation des réseaux dynamiques, où le temps y est continu.
- Une définition de la modularité valide sur ce nouvel objet.

Ouvertures

- Utiliser ce nouveau modèle pour des données de différentes natures.
- La structure des communautés est-elle indifférente du jeu de données étudié?