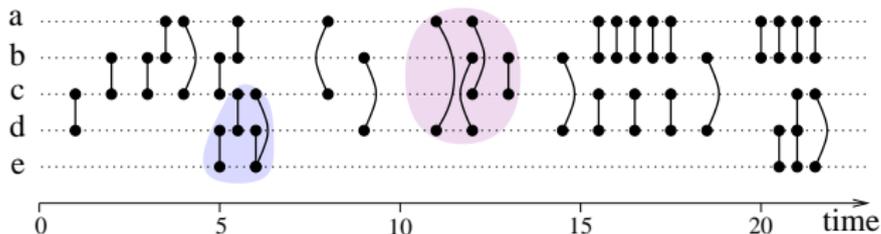


# Les flots de liens pour la modélisation des interactions temporelles



Clémence Magnien

<http://complexnetworks.fr>

[clemence.magnien@lip6.fr](mailto:clemence.magnien@lip6.fr)

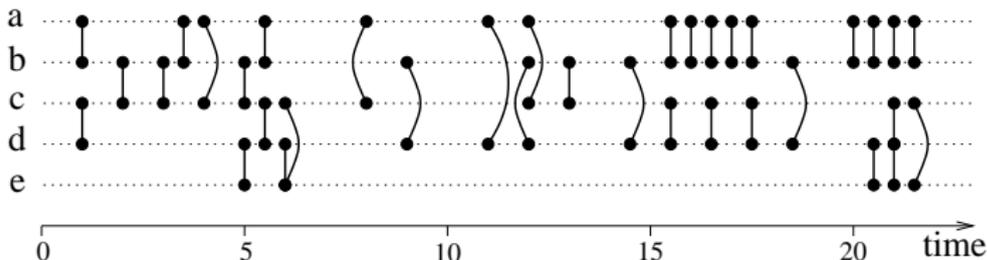
LIP6 – CNRS and UPMC

# Interactions temporelles

## De nombreux exemples

échanges d'email, trafic réseau, paiements,  
contacts physiques, appels téléphoniques, navigation Web,

...



$$L = (t_i, U_i, V_i)_{i=1..n}$$

$t_i \in [\alpha, \omega]$  : temps

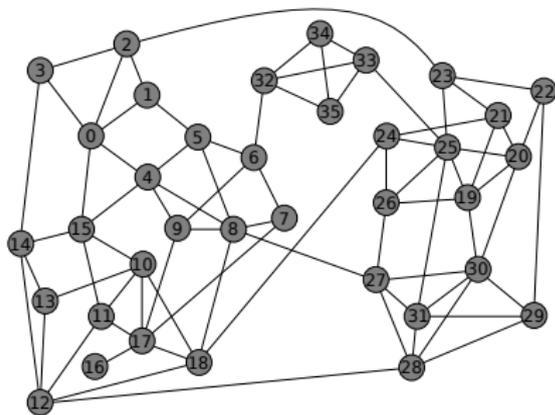
$U_i, V_i$  : nœuds

↪ Déjà beaucoup étudié

# Situation actuelle (1/3)

**approche centrée sur les liens :  $\{(a, b)\}$**

*relations, structure*

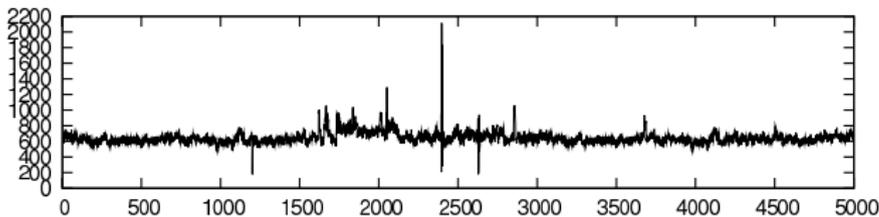


↪ **théorie des graphes / science des réseaux**  
densité, degrés, clustering,  
centralités, chemins,  
diamètre, distances, etc.

## Situation actuelle (2/3)

**approche centrée sur le temps :  $\{(t, f(t))\}$**

*événements, séries temporelles*



↪ **traitement du signal / théorie des événements  
discrets**

fréquence, vitesse, temps inter-événements,  
accélération, auto-similarité, périodicité, etc.

## Situation actuelle (3/3)

### **suites de graphes**

*temps découpé en tranches  
un graphe par tranche*

### **time-varying graphs (TVG)**

*graphe avec des étiquettes sur les liens  
étiquettes = temps d'existence*

↪ **améliore les approches graphe et signal**  
beaucoup de problèmes

# Remarque sur les différents modèles

## Formellement

Même pouvoir de représentation

## Intuition

Deux conceptions possibles

- graphe qui évolue (plus ou moins brutalement)
- **suite temporelles de liens**

**Remarque** : lié au rapport entre les échelles de temps

- du réseau
- des chemins ?

## Notre proposition

**un langage pour les flots de  
liens  
comme la théorie des graphes  
pour les réseaux**

traiter *directement* des flots de liens

les décrire, les modéliser, etc.

utiliser la richesse de leur nature structure+temps

↔ densité, coefficients de *clustering*, cliques,  
centralités, chemins, communautés, etc

Flots de liens

Clémence  
Magnien

complexnetworks.fr

Préliminaires

Densité

Chemins  
temporels

Communautés

Importance  
des nœuds

Préliminaires

Définitions et jeux de  
données

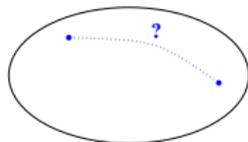
Results

Conclusion

# Densité et notions proches

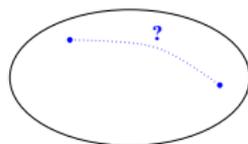
# Densité dans les graphes

Densité dans un graphe :  
proba. que deux nœuds  
aléatoires soient liés

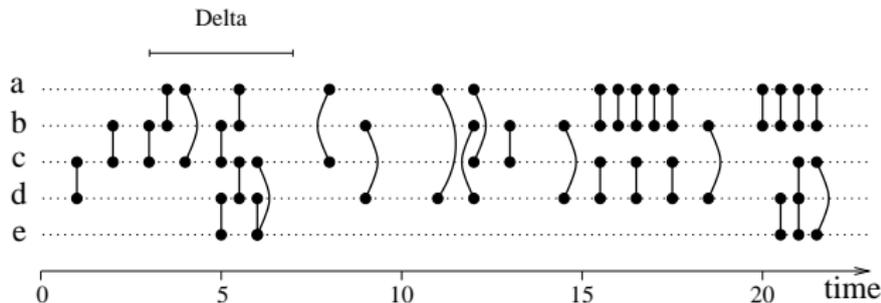


$\Delta$ -densité

Densité dans un graphe :  
proba. que deux nœuds  
aléatoires soient liés



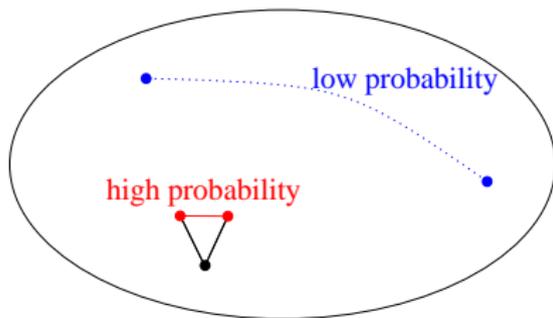
**Flots de liens** : étant donné  $\Delta$ ,  
**proba. que deux nœuds aléatoires soient liés**  
***pendant un intervalle de durée  $\Delta$  aléatoire***



Rmq :  $\Delta = \omega - \alpha \longrightarrow$  densité du graphe induit

# Coefficient de *clustering* dans les graphes

intuition : “mes amis sont amis entre eux”  
densité globale faible  
densité locale forte



**coefficient de clustering :**  
**densité du voisinage**

*à quel point tous les voisins  
sont liés*

# Coefficient de *clustering* dans les flots de liens

la même chose !

$\Delta$ -densité du voisinage

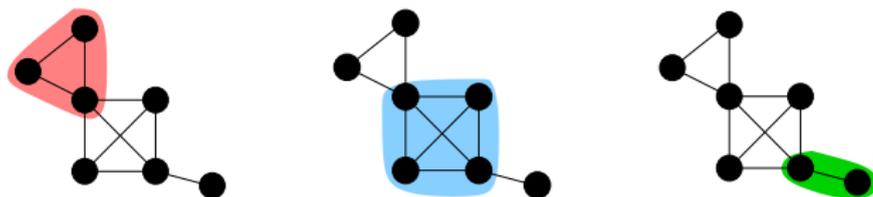
*à quel point tous les voisins  
interagissent tout le temps*

**bien plus grand que la  $\Delta$ -densité globale**

# Cliques (maximales) dans les graphes

Clique : sous-graphe de densité 1

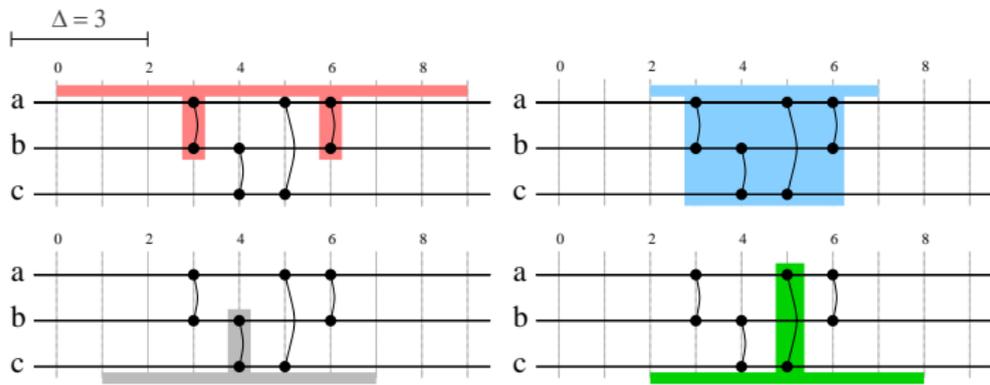
*tous les nœuds sont liés*



# Cliques (maximales) dans les flots de liens

**la même chose** : sous-flot (maximal) de  $\Delta$ -densité 1

*tous les nœuds interagissent au moins tous les  $\Delta$*



Flots de liens

Clémence  
Magnien

complexnetworks.fr

Préliminaires

Densité

**Chemins  
temporels**

Communautés

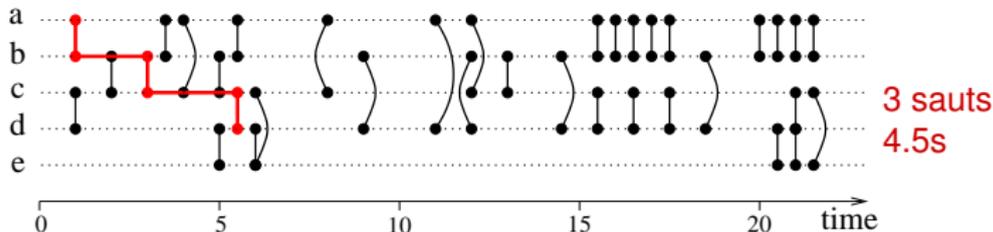
Importance  
des nœuds

Préliminaires  
Définitions et jeux de  
données  
Results

Conclusion

# Chemins temporels

## Chemins temporels



↔ longueur vs durée

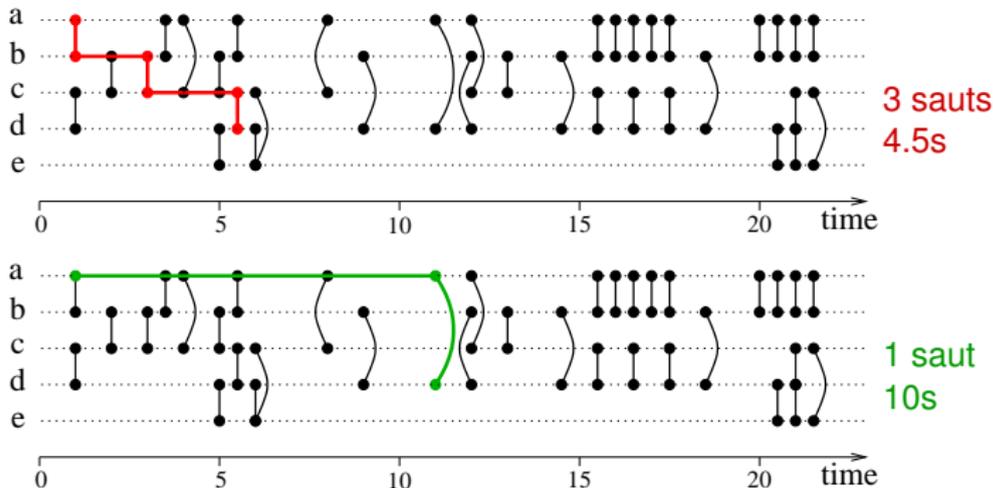
↔ distance ? **plus rapide** vs **moins de sauts**

↔ arbres, inondations

↔ (forte) connectivité

déjà "classique" dans les réseaux mobiles et phénomènes de diffusion

## Chemins temporels



↪ longueur vs durée

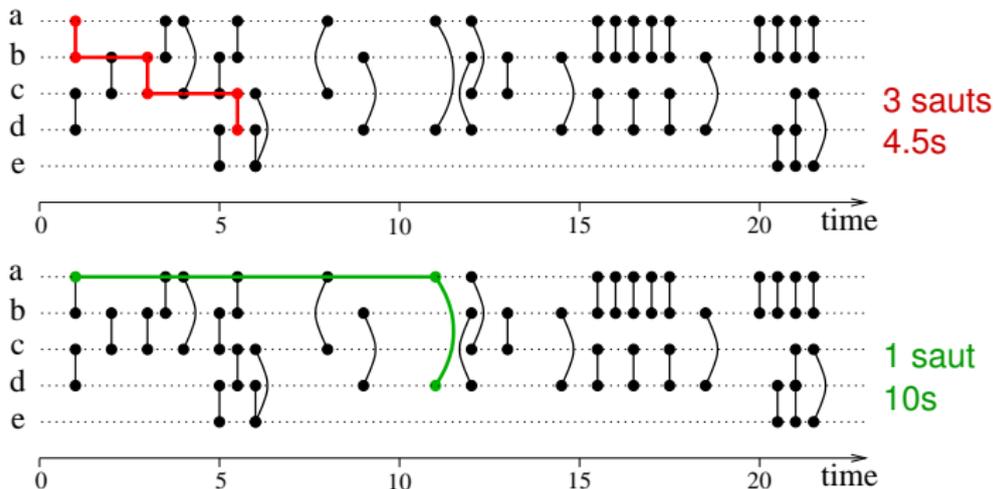
↪ distance ? **plus rapide** vs **moins de sauts**

↪ arbres, inondations

↪ (forte) connexité

déjà "classique" dans les réseaux mobiles et phénomènes de diffusion

## Chemins temporels



↪ longueur vs durée

↪ distance ? **plus rapide** vs **moins de sauts**

↪ arbres, inondations

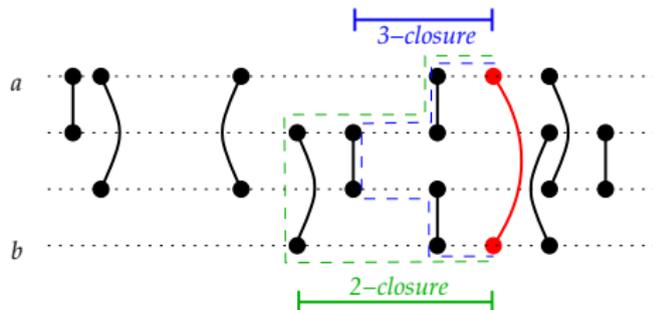
↪ (forte) connexité

déjà "classique" dans les réseaux mobiles et phénomènes de diffusion

# $k$ -clôture

$k$ -clôture de  $(t, a, b)$  :

**temps pour que  $a$  et  $b$  soient à distance  $\leq k$**



Remarques :

$\neq$  chemins temporels

$k = 1 \rightarrow$  temps inter-contacts

$k = 2 \rightarrow$  coefficient de *clustering*

*mélange de temps et de structure*

Flots de liens

Clémence  
Magnien

complexnetworks.fr

Préliminaires

Densité

Chemins  
temporels

**Communautés**

Importance  
des nœuds

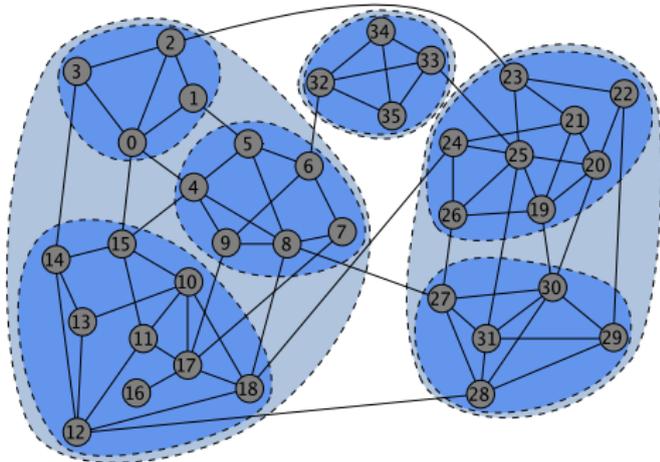
Préliminaires  
Définitions et jeux de  
données  
Results

Conclusion

# Communautés

# Communautés dans les graphes

## sous-graphes denses faiblement reliés

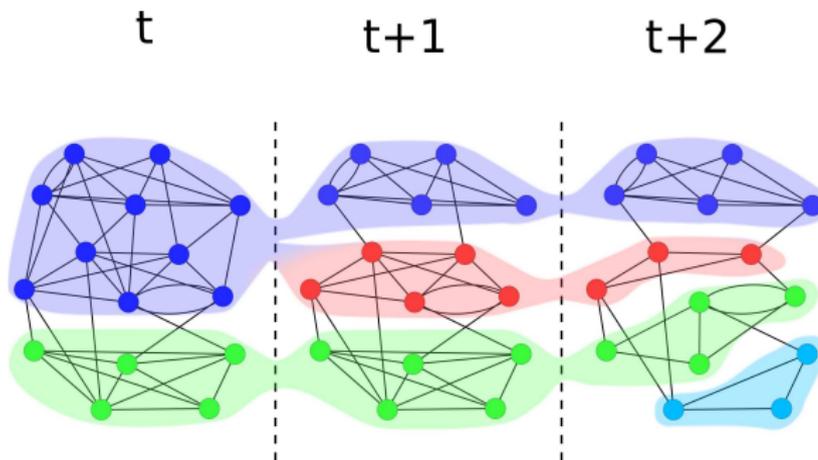


ex : groupes d'amis, d'ordinateurs, de produits, ...

comment les définir ? les détecter ?  
hiérarchies ? recouvrements ? ...

# Communautés dans les graphes *dynamiques*

## évolution des communautés dans un graphe

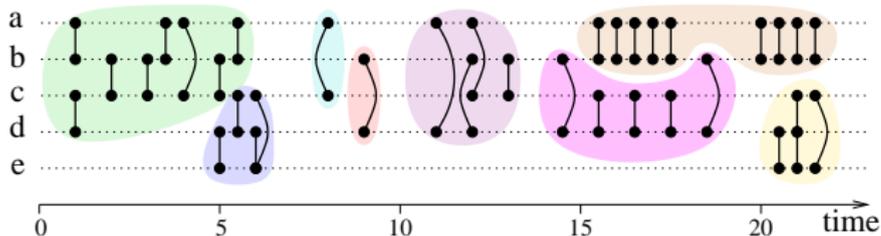


ex : groupes d'amis évoluant au fil du temps

# Communautés dans les flots de liens

## sous-flots denses faiblement reliés

i.e. des séries d'interactions temporellement et structurellement denses



ex : discussions, réunions, sessions, ...

**flots de liens  $\neq$  graphes dynamiques**

*En cours :*  
déttection (fonctions de qualité)  
vérité de terrain (fil dans une liste de discussion)

# Outline

Flots de liens

Clémence  
Magnien

complexnetworks.fr

Préliminaires

Densité

Chemins  
temporels

Communautés

**Importance  
des nœuds**

Préliminaires  
Définitions et jeux de  
données  
Results

Conclusion

- 1 Préliminaires
- 2 Densité
- 3 Chemins temporels
- 4 Communautés
- 5 Importance des nœuds**
  - Préliminaires
  - Définitions et jeux de données
  - Results
- 6 Conclusion

# Question : importance d'un nœud ?

Souvent liée à la **longueur des chemins** :  
rôle dans la diffusion (messages, rumeurs, maladies, ...)

Notions introduites pour les graphes :

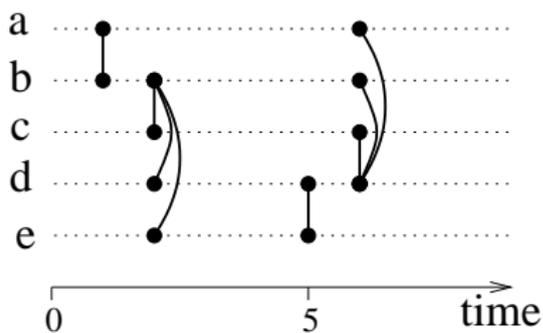
## Centralité de proximité

Somme des inverses des distances aux autres nœuds  
**grand = proche des autres nœuds**

## Centralité d'intermédiation

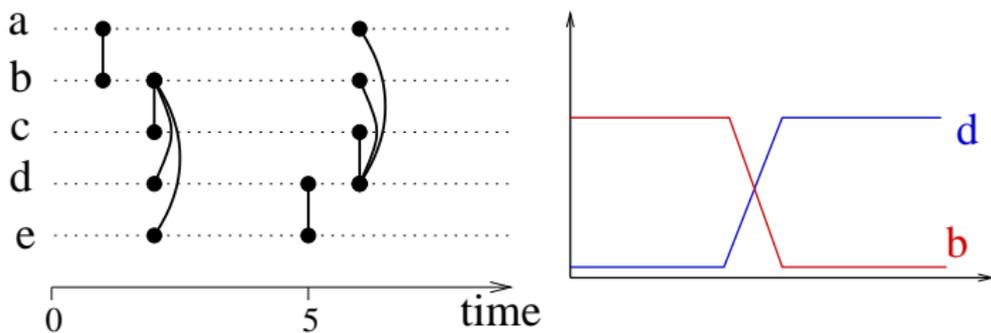
**Flots de liens ?**

## Les chemins changent au fil du temps



# Notre positionnement

Les chemins changent au fil du temps



L'importance des nœuds change

# Travaux existants dans le cas dynamique

Des extensions de la centralité ont été proposées

Soit

- Les chemins commencent au début des données  
Ignore les autres chemins
- Considèrent tous les chemins

**Valeur unique** (importance globale sur toute la durée)

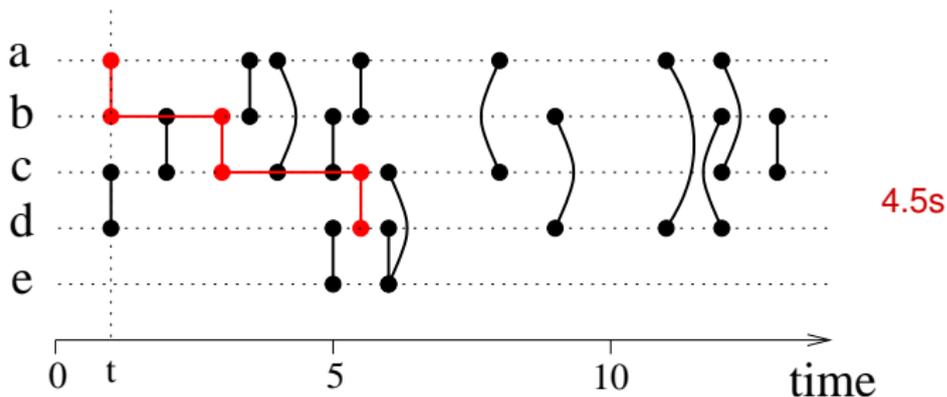
Besoin d'étudier **l'évolution** de l'importance

## Chemin temporel

Chemin temporel de  $A$  à  $B$  partant au temps  $t$  $(A, x_1, t_1), (x_1, x_2, t_2), \dots, (x_n, B, t_n)$ 

- $t_i < t_{i+1}$
- $t_1 \geq t$

Plusieurs variantes

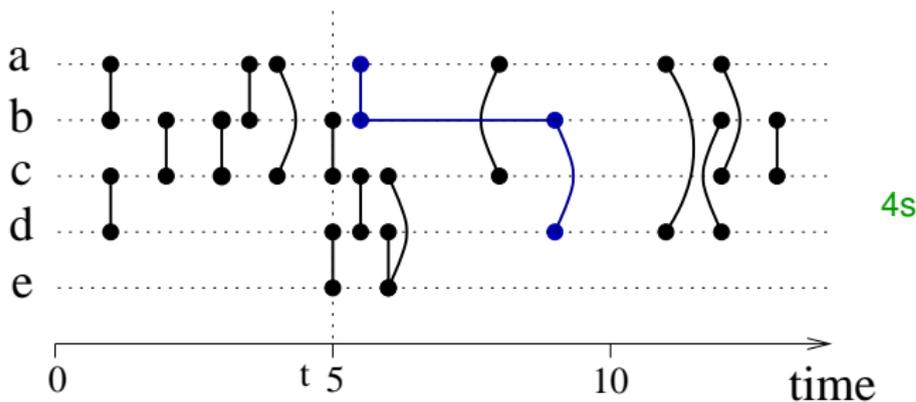


## Chemin temporel

Chemin temporel de  $A$  à  $B$  partant au temps  $t$  $(A, x_1, t_1), (x_1, x_2, t_2), \dots, (x_n, B, t_n)$ 

- $t_i < t_{i+1}$
- $t_1 \geq t$

Plusieurs variantes



# Proximité temporelle

## Distance temporelle au temps $t$ :

- $d_t(u, v) = t_a - t$ ,

$t_a$  : plus petit temps d'arrivée d'un chemin partant à  $t$

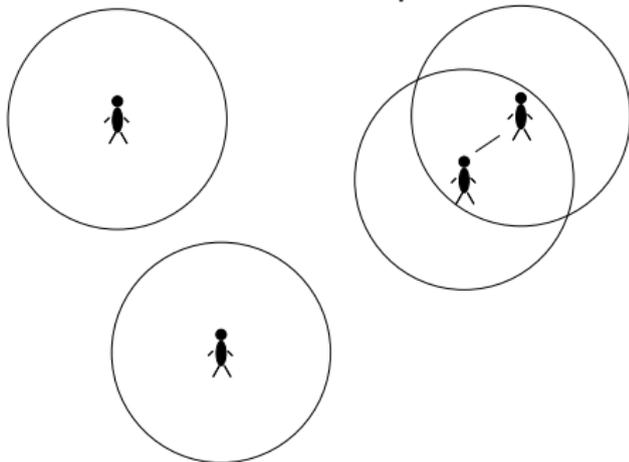
## Proximité temporelle au temps $t$

$$C_t(u) = \sum_{v \neq u} \frac{1}{d_t(u, v)}$$

# Données (1) – Rollernet

[Tournoux *et al*, 2009]

Proximité physiques entre personnes  
Mesurée avec des capteurs



Rando roller

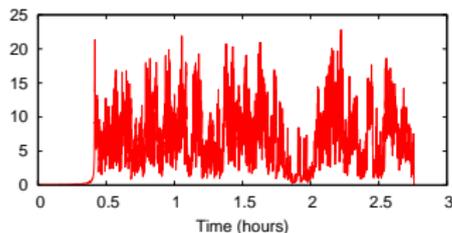
- 62 nœuds
- ~ 3h

# Données (2) – email, Enron

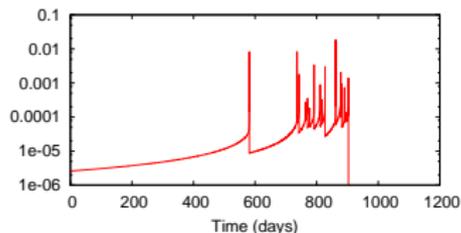
## Emails entre les employés d'Enron

- 151 employés
- plus de trois ans

# Proximité de nœuds aléatoires



Rollernet



Enron

## Problèmes

- Fluctue beaucoup
- Difficile à interpréter
- Comparaison avec d'autres nœuds

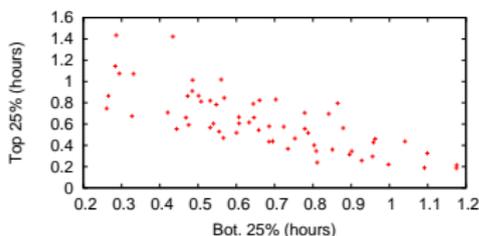
## Comparison globale

- Proximité pour tous les nœuds à tous les instants
- → rang des nœuds à chaque instant

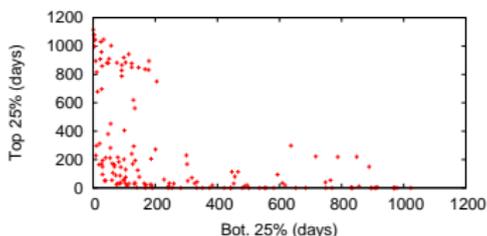
# Statistiques globales

## Pour chaque nœud

Temps passé dans les 25% rangs plus élevés/faibles



Rollernet



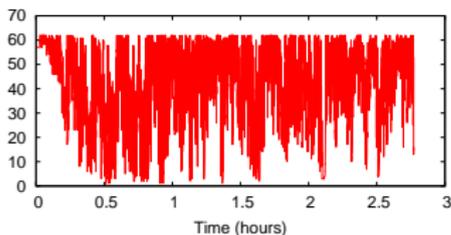
Enron

## Observations

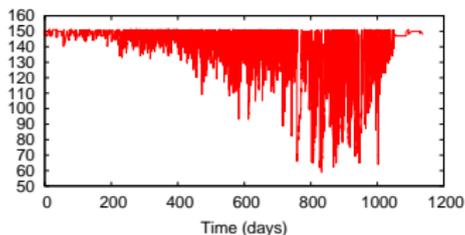
- Rollernet : aucun nœud n'est important pendant plus de la moitié du temps
- Enron : quelques nœuds (in-)importants globalement

## Exemple

Rang des nœuds les plus longtemps dans les 25% plus élevés



Rollernet



Enron

## Observations

- Plus important  $\neq$  globalement important
- Globalement important  $\neq$  toujours important

## Résultats

- Importance varie au fil du temps
- Notion d'importance globale pas toujours pertinente
- Observations dépendantes du jeu de données

## Perspectives

- Prédiction
- Importance des liens
- Détection d'anomalies

# Outline

Flots de liens

Clémence  
Magnien

complexnetworks.fr

Préliminaires

Densité

Chemins  
temporels

Communautés

Importance  
des nœuds

Préliminaires  
Définitions et jeux de  
données  
Results

Conclusion

- 1 Préliminaires
- 2 Densité
- 3 Chemins temporels
- 4 Communautés
- 5 Importance des nœuds
  - Préliminaires
  - Définitions et jeux de données
  - Results
- 6 Conclusion

## les flots de liens modélisent les interactions temporelles flots de liens $\neq$ graphes dynamiques

- **Fait** : densité, algorithmique des cliques, chemins, coefficients de *clustering*, communautés.
- **En cours** : communautés en pratique, détection d'événements et de communautés, centralités, liens avec les TVG
- **Cas d'étude** : listes de discussion (Debian), appels téléphoniques (D4D), trafic réseau (Mawi, entreprises), mobilité/contacts (crowdad, sociopatterns), transactions financières (bitcoins, achats en ligne), etc.
- **Extensions** : force, durée, orientation, etc. des interactions  $\rightarrow$  flots de liens pondérés, orientés, bipartis, etc.