

# Diffusion dans les réseaux dynamiques

Jérémie Dumas  
Maître de Stage : Christophe Crespelle

Laboratoire d'Informatique de Paris 6  
4 place Jussieu  
75005 Paris

Juin-Juillet 2010



# Sommaire

## Généralités

Définitions sur les graphes dynamiques

## Trajets

Où l'on parle de chemins

## Contacts

Où l'on compte des liaisons entre les gens

## Flots

Dans un graphe dynamique

## Courbes

# Introduction

## Le contexte

- ▶ Graphe dynamique : évolue au cours du temps
- ▶ Sommets fixes  $\rightarrow V$
- ▶ Arêtes dynamiques : ajout, suppression  $\rightarrow E_t$
- ▶ Diffusion à partir d'un ensemble  $S_0 \subseteq V$

# Introduction

## Le contexte

- ▶ Graphe dynamique : évolue au cours du temps
- ▶ Sommets fixes  $\rightarrow V$
- ▶ Arêtes dynamiques : ajout, suppression  $\rightarrow E_t$
- ▶ Diffusion à partir d'un ensemble  $S_0 \subseteq V$

# Introduction

## Le contexte

- ▶ Graphe dynamique : évolue au cours du temps
- ▶ Sommets fixes  $\rightarrow V$
- ▶ Arêtes dynamiques : ajout, suppression  $\rightarrow E_t$
- ▶ Diffusion à partir d'un ensemble  $S_0 \subseteq V$

# Introduction

## Le contexte

- ▶ Graphe dynamique : évolue au cours du temps
- ▶ Sommets fixes  $\rightarrow V$
- ▶ Arêtes dynamiques : ajout, suppression  $\rightarrow E_t$
- ▶ Diffusion à partir d'un ensemble  $S_0 \subseteq V$

# Données utilisées

## Les expériences

- ▶ I-Bird (Hôpital) : 800 pers., 6 mois. (Non exploité.)
- ▶ Infocom (Conférence) : 40 pers., 3 jours.
- ▶ Reality Mining (MIT) : 100 pers., 9 mois.

# Données utilisées

## Les expériences

- ▶ I-Bird (Hôpital) : 800 pers., 6 mois. (Non exploité.)
- ▶ Infocom (Conférence) : 40 pers., 3 jours.
- ▶ Reality Mining (MIT) : 100 pers., 9 mois.



# Données utilisées

## Les expériences

- ▶ I-Bird (Hôpital) : 800 pers., 6 mois. (Non exploité.)
- ▶ Infocom (Conférence) : 40 pers., 3 jours.
- ▶ Reality Mining (MIT) : 100 pers., 9 mois.

# Graphe dynamique

## Définition

- ▶ Graphe dynamique : séquence  $(G_t = (V_t, E_t))_{0 \leq t < t_m}$
- ▶ Voisinage de  $u$  dans  $G_t$  :  $\mathcal{N}_t(u)$
- ▶ Chemin dynamique (trajet) : séquence  $c_{t_0} = u_0, \dots, u_p$  telle que  $\forall i \in \llbracket 1, p \rrbracket, u_{i-1}u_i \in E_{i-1+t_0} \vee u_{i-1} = u_i$

# Graphe dynamique

## Définition

- ▶ Graphe dynamique : séquence  $(G_t = (V_t, E_t))_{0 \leq t < t_m}$
- ▶ Voisinage de  $u$  dans  $G_t$  :  $\mathcal{N}_t(u)$
- ▶ Chemin dynamique (trajet) : séquence  $c_{t_0} = u_0, \dots, u_p$  telle que  $\forall i \in \llbracket 1, p \rrbracket, u_{i-1}u_i \in E_{i-1+t_0} \vee u_{i-1} = u_i$

# Graphe dynamique

## Définition

- ▶ Graphe dynamique : séquence  $(G_t = (V_t, E_t))_{0 \leq t < t_m}$
- ▶ Voisinage de  $u$  dans  $G_t$  :  $\mathcal{N}_t(u)$
- ▶ Chemin dynamique (trajet) : séquence  $c_{t_0} = u_0, \dots, u_p$  telle que  $\forall i \in \llbracket 1, p \rrbracket, u_{i-1}u_i \in E_{i-1+t_0} \vee u_{i-1} = u_i$

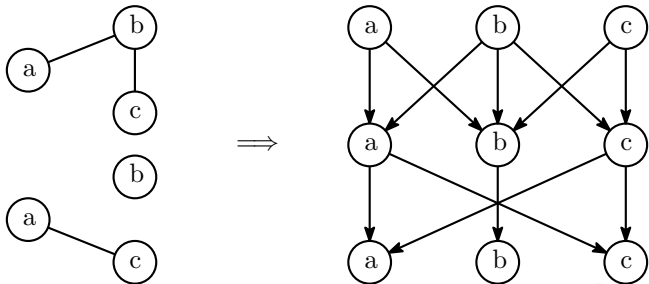
# Graphe de transition

## Définition

C'est un graphe unique qui représente la séquence  $G_t$ .

Notons  $\mathcal{G} = (V^T, E^T)$  le graphe **orienté** tel que :

- ▶  $V^T = \{(x, t), x \in V, 0 \leq t \leq t_m\}$
- ▶  $E^T = \{(x, t)(y, t+1), x = y \vee (x, y) \in E_t, 0 \leq t < t_m\}$



# Temps de diffusion

Et nombre de trajets

## Définition

- ▶  $\mathcal{T}_{S_0}(v)(t) = \min \left\{ d(c), \begin{array}{l} c \text{ trajets de } S_0 \text{ à } v \\ \text{de date d'arrivée } t_a(c) = t \end{array} \right\}$
- ▶  $\mathcal{C}_{S_0}(v)(t) = |\{c, d(c) = \mathcal{T}_{S_0}(v)(t) \wedge t_a(c) = t\}|$ ,  
nombre de tels trajets

## Complexité

Calcul de  $\mathcal{T}_{S_0}$  et  $\mathcal{C}_{S_0}$  pour tout  $v, t : \mathcal{O}((n + m_{max})p)$

# Temps de diffusion

## Et nombre de trajets

### Définition

- ▶  $\mathcal{T}_{S_0}(v)(t) = \min \left\{ d(c), \begin{array}{l} c \text{ trajets de } S_0 \text{ à } v \\ \text{de date d'arrivée } t_a(c) = t \end{array} \right\}$
- ▶  $\mathcal{C}_{S_0}(v)(t) = |\{c, d(c) = \mathcal{T}_{S_0}(v)(t) \wedge t_a(c) = t\}|$ ,  
nombre de tels trajets

### Complexité

Calcul de  $\mathcal{T}_{S_0}$  et  $\mathcal{C}_{S_0}$  pour tout  $v, t : \mathcal{O}((n + m_{max})p)$

# Temps de diffusion

Et nombre de trajets

## Définition

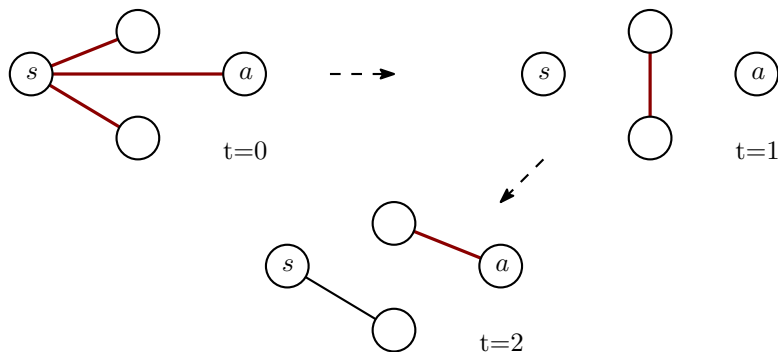
- ▶  $\mathcal{T}_{S^0}(v)(t) = \min \left\{ d(c), \begin{array}{l} c \text{ trajets de } S_0 \text{ à } v \\ \text{de date d'arrivée } t_a(c) = t \end{array} \right\}$
- ▶  $\mathcal{C}_{S^0}(v)(t) = |\{c, d(c) = \mathcal{T}_{S^0}(v)(t) \wedge t_a(c) = t\}|$ ,  
nombre de tels trajets

## Complexité

Calcul de  $\mathcal{T}_{S^0}$  et  $\mathcal{C}_{S^0}$  pour tout  $v, t$  :  $\mathcal{O}((n + m_{max})p)$

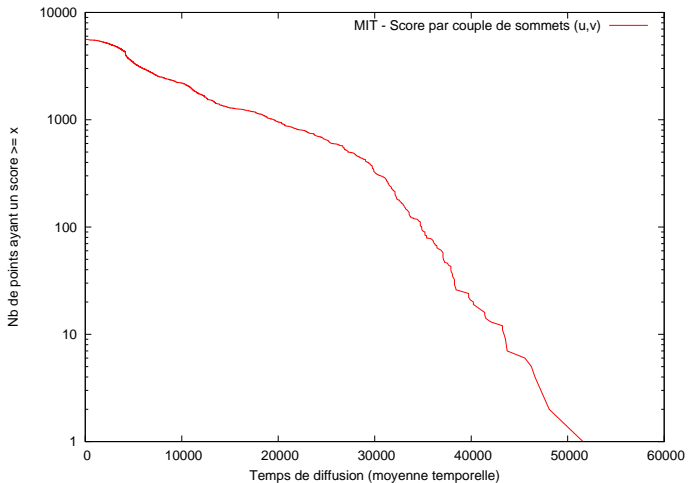


## Exemple

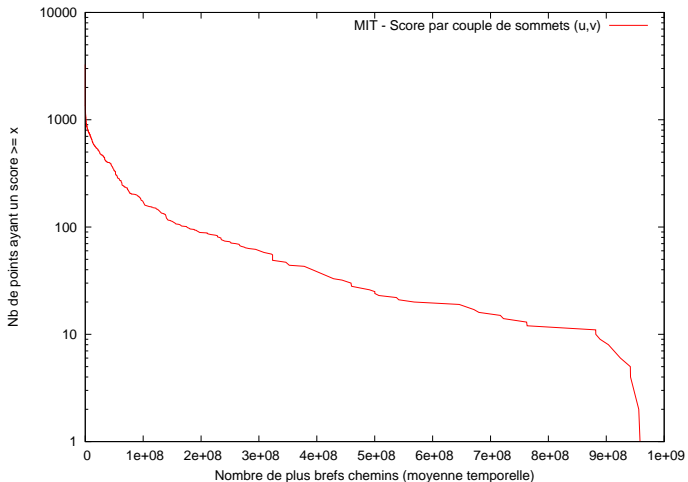


Temps de diffusion	$\mathcal{T}_{\{s\}}(a)(3) = 2$
Nb de trajets	$\mathcal{C}_{\{s\}}(a)(3) = 3$

# Distributions cumulatives inverses



# Distributions cumulatives inverses



# Nombre de contacts

## Contacts simples

### Idée

Compter le nb de contact entre  $S_0$  et  $v$  sur les  $\gamma$  derniers graphes.

### Définition (Indice de contact)

$$I_{S_0, \gamma}(v)(t) = \sum_{i=t-\gamma+1}^t |S^0 \cap \mathcal{N}_i(v)|$$

### Complexité

Calcul de  $\mathcal{T}_{S_0}$  et  $\mathcal{C}_{S_0}$  pour tout  $v, t$  :  $\mathcal{O}((n + m_{max})p \times \gamma)$

# Nombre de contacts

## Contacts simples

### Idée

Compter le nb de contact entre  $S_0$  et  $v$  sur les  $\gamma$  derniers graphes.

### Définition (Indice de contact)

$$I_{S^0, \gamma}(v)(t) = \sum_{i=t-\gamma+1}^t |S^0 \cap \mathcal{N}_i(v)|$$

### Complexité

Calcul de  $\mathcal{T}_{S^0}$  et  $\mathcal{C}_{S^0}$  pour tout  $v, t$  :  $\mathcal{O}((n + m_{max})p \times \gamma)$

# Nombre de contacts

## Contacts simples

### Idée

Compter le nb de contact entre  $S_0$  et  $v$  sur les  $\gamma$  derniers graphes.

### Définition (Indice de contact)

$$I_{S^0, \gamma}(v)(t) = \sum_{i=t-\gamma+1}^t |S^0 \cap \mathcal{N}_i(v)|$$

### Complexité

Calcul de  $\mathcal{T}_{S^0}$  et  $\mathcal{C}_{S^0}$  pour tout  $v, t$  :  $\mathcal{O}((n + m_{max})p \times \gamma)$

# Nombre de contacts

## Contacts étendus

### Diffusion

On fait évoluer  $S_0$  à partir de l'instant  $t_d$  :  $t \rightarrow S_t = f_c(S^0, t_d, t)$ .

- ▶ On ajoute à  $S_t$  ses voisins dans  $G_t$ .
- ▶ Même complexité de calcul que  $I_{S^0, \gamma}(v)(t)$

### Définition (Indice de contact étendu)

$$I_{S^0, \gamma}^*(v)(t) = \sum_{i=t-\gamma+1}^t |f_c(S^0, t-\gamma+1, i) \cap \mathcal{N}_i(v)|$$

# Nombre de contacts

## Contacts étendus

### Diffusion

On fait évoluer  $S_0$  à partir de l'instant  $t_d$  :  $t \rightarrow S_t = f_c(S^0, t_d, t)$ .

- ▶ On ajoute à  $S_t$  ses voisins dans  $G_t$ .
- ▶ Même complexité de calcul que  $I_{S^0, \gamma}(v)(t)$

### Définition (Indice de contact étendu)

$$I_{S^0, \gamma}^*(v)(t) = \sum_{i=t-\gamma+1}^t |f_c(S^0, t-\gamma+1, i) \cap \mathcal{N}_i(v)|$$



# Nombre de contacts

## Contacts étendus

### Diffusion

On fait évoluer  $S_0$  à partir de l'instant  $t_d$  :  $t \rightarrow S_t = f_c(S^0, t_d, t)$ .

- ▶ On ajoute à  $S_t$  ses voisins dans  $G_t$ .
- ▶ Même complexité de calcul que  $I_{S^0, \gamma}(v)(t)$

### Définition (Indice de contact étendu)

$$I_{S^0, \gamma}^*(v)(t) = \sum_{i=t-\gamma+1}^t |f_c(S^0, t-\gamma+1, i) \cap \mathcal{N}_i(v)|$$

# Nombre de contacts

## Contacts étendus

### Diffusion

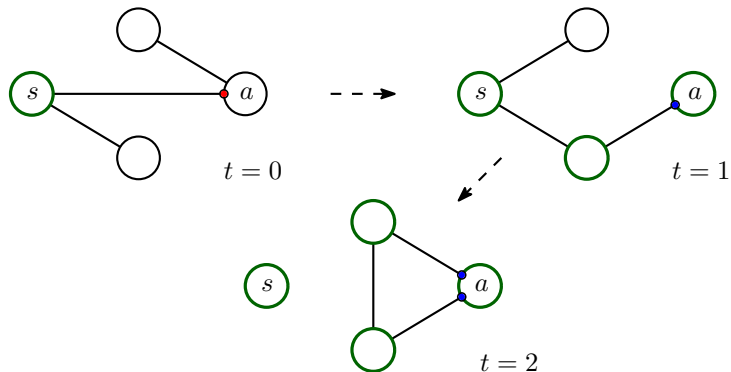
On fait évoluer  $S_0$  à partir de l'instant  $t_d : t \rightarrow S_t = f_c(S^0, t_d, t)$ .

- ▶ On ajoute à  $S_t$  ses voisins dans  $G_t$ .
- ▶ Même complexité de calcul que  $I_{S^0, \gamma}(v)(t)$

### Définition (Indice de contact étendu)

$$I_{S^0, \gamma}^*(v)(t) = \sum_{i=t-\gamma+1}^t |f_c(S^0, t-\gamma+1, i) \cap \mathcal{N}_i(v)|$$

## Exemple



Contacts simples	$I_{\{s\},\gamma}(a)(2) = 1$
Contacts étendus	$I_{\{s\},\gamma}^*(a)(2) = 4$

# Flot dynamique

## Définition

Soit  $(G_t)_t$  un graphe dynamique pondéré + capacités  $(c_t)_t$ .

Soit une **source**  $(s, t_0)$  et un **puits**  $(a, t_1)$ .

On définit  $(\varphi_t)_{t_0 \leq t \leq t_1}$  flot dynamique s'il vérifie les contraintes :

1. **Positivité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \geq 0$
2. **Contrainte de capacité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \leq c_t(x, y)$
3. **Conservation** :  
$$\forall t_0 \leq t < t_1, \forall x \in V, \sum_y \varphi_t(y, x) = \sum_y \varphi_{t+1}(x, y)$$
4. **Conditions aux limites** :
  - 4.1  $\forall x, y \in V, x \neq s \Rightarrow \varphi_{t_0}(x, y) = 0$
  - 4.2  $\forall x, y \in V, y \neq a \Rightarrow \varphi_{t_1}(x, y) = 0$

# Flot dynamique

## Définition

Soit  $(G_t)_t$  un graphe dynamique pondéré + capacités  $(c_t)_t$ .

Soit une source  $(s, t_0)$  et un puits  $(a, t_1)$ .

On définit  $(\varphi_t)_{t_0 \leq t \leq t_1}$  flot dynamique s'il vérifie les contraintes :

1. **Positivité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \geq 0$
2. **Contrainte de capacité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \leq c_t(x, y)$
3. **Conservation** :  
$$\forall t_0 \leq t < t_1, \forall x \in V, \sum_y \varphi_t(y, x) = \sum_y \varphi_{t+1}(x, y)$$
4. **Conditions aux limites** :
  - 4.1  $\forall x, y \in V, x \neq s \Rightarrow \varphi_{t_0}(x, y) = 0$
  - 4.2  $\forall x, y \in V, y \neq a \Rightarrow \varphi_{t_1}(x, y) = 0$

# Flot dynamique

## Définition

Soit  $(G_t)_t$  un graphe dynamique pondéré + capacités  $(c_t)_t$ .

Soit une source  $(s, t_0)$  et un puits  $(a, t_1)$ .

On définit  $(\varphi_t)_{t_0 \leq t \leq t_1}$  flot dynamique s'il vérifie les contraintes :

1. **Positivité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \geq 0$
2. **Contrainte de capacité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \leq c_t(x, y)$

3. **Conservation** :

$$\forall t_0 \leq t < t_1, \forall x \in V, \sum_y \varphi_t(y, x) = \sum_y \varphi_{t+1}(x, y)$$

4. **Conditions aux limites** :

$$4.1 \quad \forall x, y \in V, x \neq s \Rightarrow \varphi_{t_0}(x, y) = 0$$

$$4.2 \quad \forall x, y \in V, y \neq a \Rightarrow \varphi_{t_1}(x, y) = 0$$

# Flot dynamique

## Définition

Soit  $(G_t)_t$  un graphe dynamique pondéré + capacités  $(c_t)_t$ .

Soit une source  $(s, t_0)$  et un puits  $(a, t_1)$ .

On définit  $(\varphi_t)_{t_0 \leq t \leq t_1}$  flot dynamique s'il vérifie les contraintes :

1. **Positivité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \geq 0$
2. **Contrainte de capacité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \leq c_t(x, y)$

3. **Conservation** :

$$\forall t_0 \leq t < t_1, \forall x \in V, \sum_y \varphi_t(y, x) = \sum_y \varphi_{t+1}(x, y)$$

4. **Conditions aux limites** :

$$4.1 \quad \forall x, y \in V, x \neq s \Rightarrow \varphi_{t_0}(x, y) = 0$$

$$4.2 \quad \forall x, y \in V, y \neq a \Rightarrow \varphi_{t_1}(x, y) = 0$$

# Flot dynamique

## Définition

Soit  $(G_t)_t$  un graphe dynamique pondéré + capacités  $(c_t)_t$ .

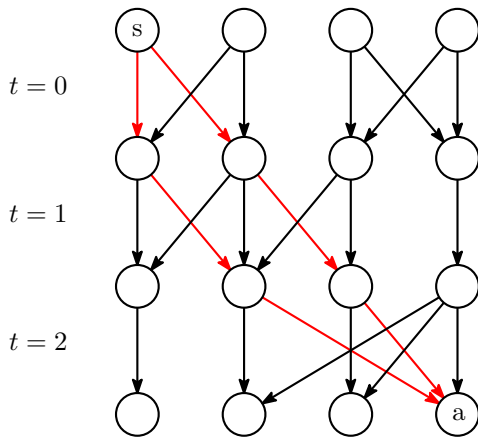
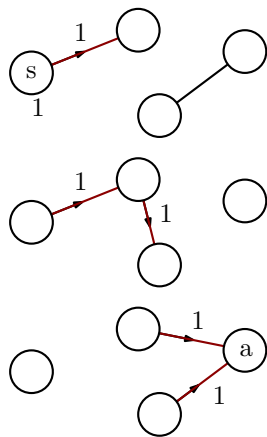
Soit une source  $(s, t_0)$  et un puits  $(a, t_1)$ .

On définit  $(\varphi_t)_{t_0 \leq t \leq t_1}$  flot dynamique s'il vérifie les contraintes :

1. **Positivité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \geq 0$
2. **Contrainte de capacité** :  $\forall t, \forall x, y \in V, \varphi_t(x, y) \leq c_t(x, y)$
3. **Conservation** :  
$$\forall t_0 \leq t < t_1, \forall x \in V, \sum_y \varphi_t(y, x) = \sum_y \varphi_{t+1}(x, y)$$
4. **Conditions aux limites** :
  - 4.1  $\forall x, y \in V, x \neq s \Rightarrow \varphi_{t_0}(x, y) = 0$
  - 4.2  $\forall x, y \in V, y \neq a \Rightarrow \varphi_{t_1}(x, y) = 0$



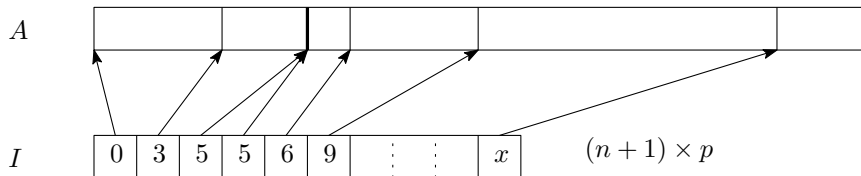
## Correspondance avec un flot usuel



Flot dynamique

$$Q_{s,\gamma}(a)(2) = 2$$

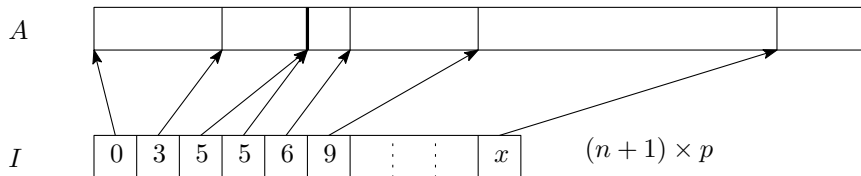
## Structure de donnée



### Remarques

- ▶ Charger le graphe de transition en mémoire : **lourd**
- ▶ Algorithme d'Edmonds-Karp sur  $\mathcal{G}$  : peu optimisé
- ▶ Mais envisageable vu la taille des données

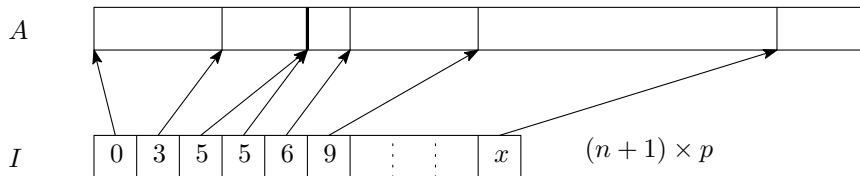
## Structure de donnée



### Remarques

- ▶ Charger le graphe de transition en mémoire : lourd
- ▶ Algorithme d'Edmonds-Karp sur  $\mathcal{G}$  : **peu optimisé**
- ▶ Mais envisageable vu la taille des données

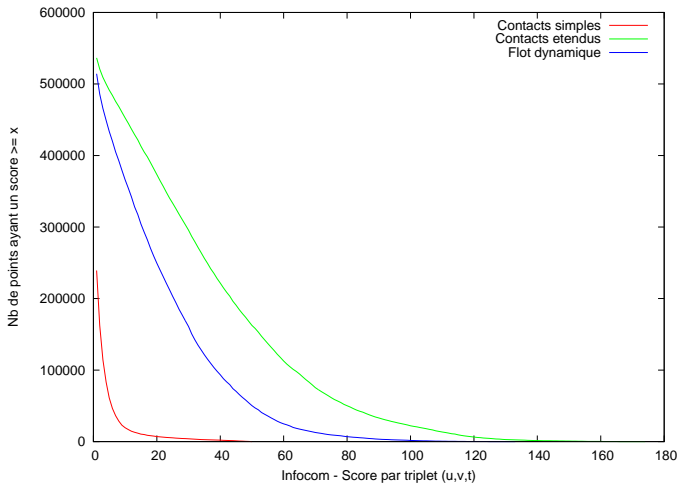
## Structure de donnée



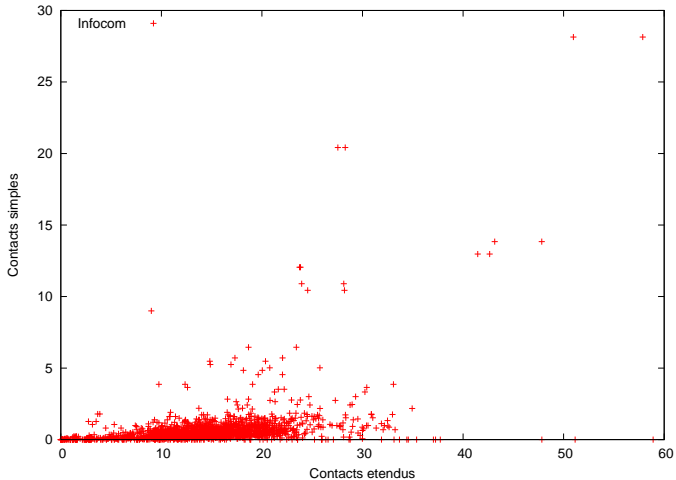
### Remarques

- ▶ Charger le graphe de transition en mémoire : lourd
- ▶ Algorithme d'Edmonds-Karp sur  $\mathcal{G}$  : peu optimisé
- ▶ Mais envisageable vu la taille des données

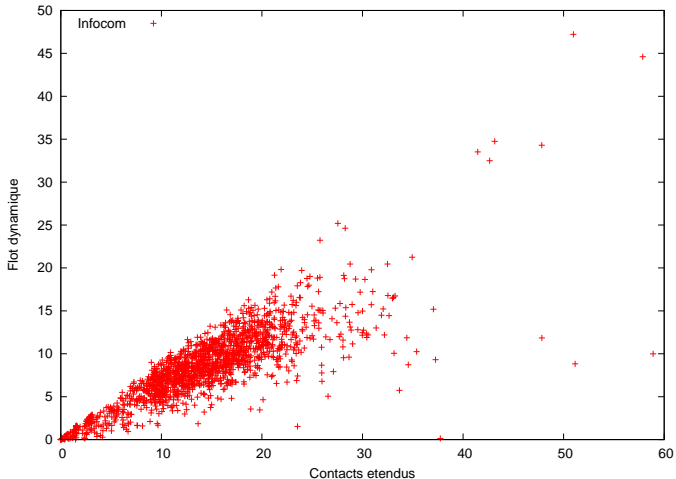
## Distribution cumulative inverse



# Corrélations



# Corrélations



# Conclusion

- ▶ Trajets les plus rapides : explosion combinatoire.
- ▶ Flot dynamique : intéressant mais encore coûteux à calculer.
- ▶ Contacts étendus : un compromis intéressant.
- ▶ Outils qui méritent d'être approfondis.



# Conclusion

- ▶ Trajets les plus rapides : explosion combinatoire.
- ▶ Flot dynamique : intéressant mais encore coûteux à calculer.
- ▶ Contacts étendus : un compromis intéressant.
- ▶ Outils qui méritent d'être approfondis.

## Conclusion

- ▶ Trajets les plus rapides : explosion combinatoire.
- ▶ Flot dynamique : intéressant mais encore coûteux à calculer.
- ▶ Contacts étendus : un compromis intéressant.
- ▶ Outils qui méritent d'être approfondis.

## Conclusion

- ▶ Trajets les plus rapides : explosion combinatoire.
- ▶ Flot dynamique : intéressant mais encore coûteux à calculer.
- ▶ Contacts étendus : un compromis intéressant.
- ▶ Outils qui méritent d'être approfondis.

Merci de votre attention.

N'hésitez pas à poser vos questions.