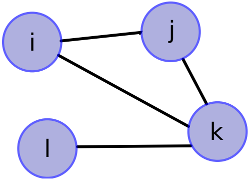
Modèle

- $G = (V, E)$
- $V = \{1, \ldots, n\}$
- Caractéristiques : $(X_i)_{1 \leq i \leq n}$



Objectif : regrouper par similarité suivant les caractéristiques

Formulation du problème

- Problème :

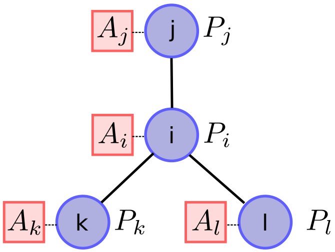
$$\min_{P \in \mathcal{P}} w(P) = \frac{1}{n^2} \sum_{1 \leq i,j \leq n} D(X_i, X_j) \Phi_P(X_i, X_j)$$

  - D : mesure de dissimilarité
  - $\mathcal{P}$: ensemble des partitions admissibles
  - Φ_P : indicatrice de *cluster*
- Contraintes supplémentaires :
  - Distribution du calcul
  - Certains $D(X_i, X_j)$ sont inaccessibles

Principe des algorithmes de *gossip*

Le nœud i :

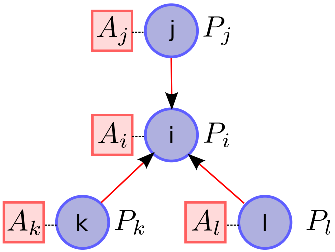
- dispose de l'information A_i
- effectue une estimation P_i



Principe des algorithmes de *gossip*

Le nœud i :

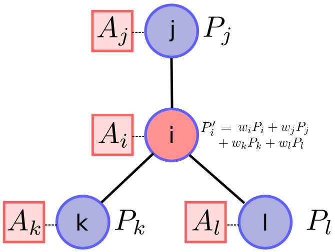
- dispose de l'information A_i
- effectue une estimation P_i
- reçoit les estimations de ses voisins



Principe des algorithmes de *gossip*

Le nœud i :

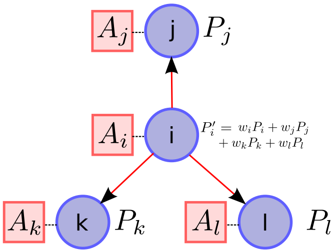
- dispose de l'information A_i
- effectue une estimation P_i
- reçoit les estimations de ses voisins



Principe des algorithmes de *gossip*

Le nœud i :

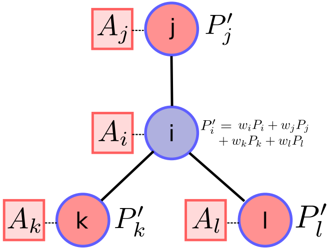
- dispose de l'information A_i
- effectue une estimation P_i
- reçoit les estimations de ses voisins
- transmet son estimation à ses voisins



Principe des algorithmes de *gossip*

Le nœud i :

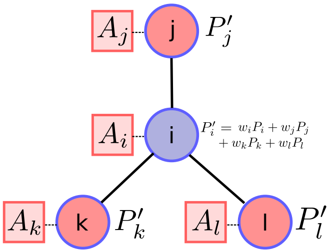
- dispose de l'information A_i
- effectue une estimation P_i
- reçoit les estimations de ses voisins
- transmet son estimation à ses voisins



Principe des algorithmes de *gossip*

Le nœud i :

- dispose de l'information A_i
- effectue une estimation P_i
- reçoit les estimations de ses voisins
- transmet son estimation à ses voisins



Nombreuses variantes possibles :

- Transmission de l'information
  - synchrone/asynchrone
  - stochastique/totale
- Limitation du volume de données transmises
- Graphe dynamique
- Données dynamiques

Exemple : K-means

K-means :

- explications k-means

K-means par *gossip*

Estimation

- Pour $P \in \mathcal{P}$, on aimerait calculer :

$$w(P) = \frac{1}{n^2} \sum_{1 \leq i,j \leq n} D(X_i, X_j) \Phi_P(X_i, X_j)$$

- Problème : certains $D(X_i, X_j)$ inaccessibles
- Idée : estimer $f : x \mapsto \mathbb{E}\left[D(x, X) \Phi_P(x, X)\right]$