



21033

Présentation et soutenance

Notions de centralité dans les graphes

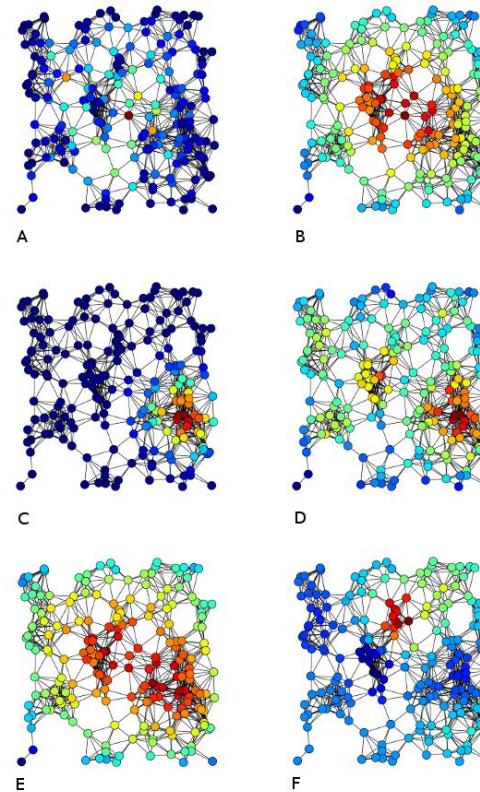
QI Zhirui
ZHU Zhihao
HUANG Bozhang

Tableau de matière

- **01** *Contexte de la centralité*
- **02** *Partie algorithmique*
- **03** *Partie implémentation*
- **04** *Partie analyse de données*
- **05** *Résumé*

01

Contexte de la centralité



Définitions de base sur les graphes



Graphe non-orienté

$G = (V; E)$

- V un ensemble de nœuds
- E un ensemble d'arêtes



Chemin

un chemin d'origine x et d'extrémité y, noté $[x; y]$



Graphe connexe

un graphe non orienté $G = (V; E)$ est dit connexe si quels que soient les nœuds u et v de G, il existe un chemin reliant u et v.

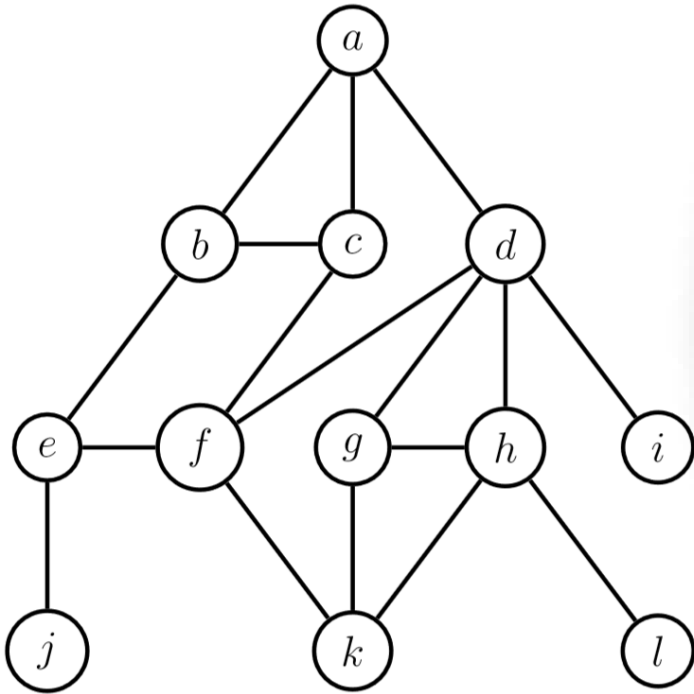


Densité d'un graphe

Ce paramètre mesure si le graphe a beaucoup d'arêtes ou peu

$$\delta = \frac{2|E|}{|V| \cdot (|V| - 1)}$$

Définitions de distance et des caractéristiques des centralités



Centralité de degré

$$k(u) = |\{v | (u, v) \in E\}|$$

Centralité de proximité

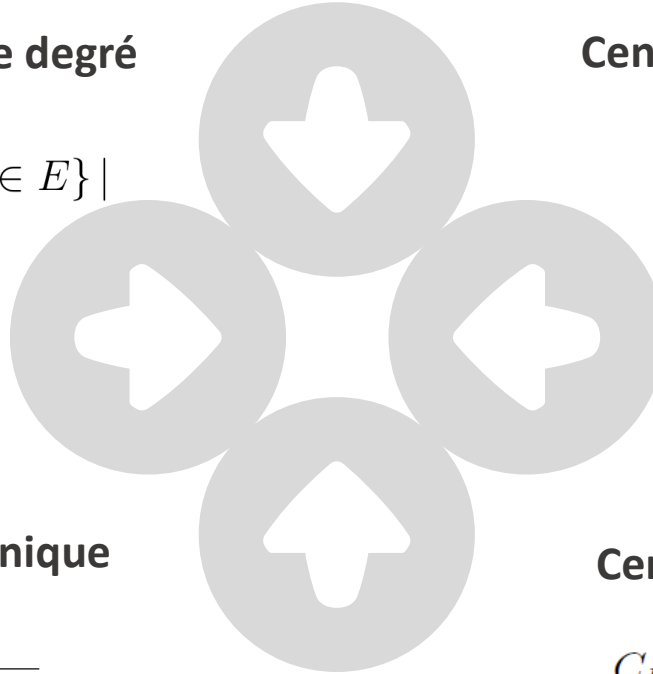
$$C_C(x) = \frac{1}{\sum_{y \in V, y \neq x} d(x, y)}$$

Centralité harmonique

$$C_H(x) = \sum_{y \neq x} \frac{1}{d(y, x)}$$

Centralité d'intermédiation

$$C_B(x) = \sum_{\substack{s \neq v, v \neq t, s \neq t \\ s, v, t \in V}} \frac{\sigma_{st}(x)}{\sigma_{st}}$$

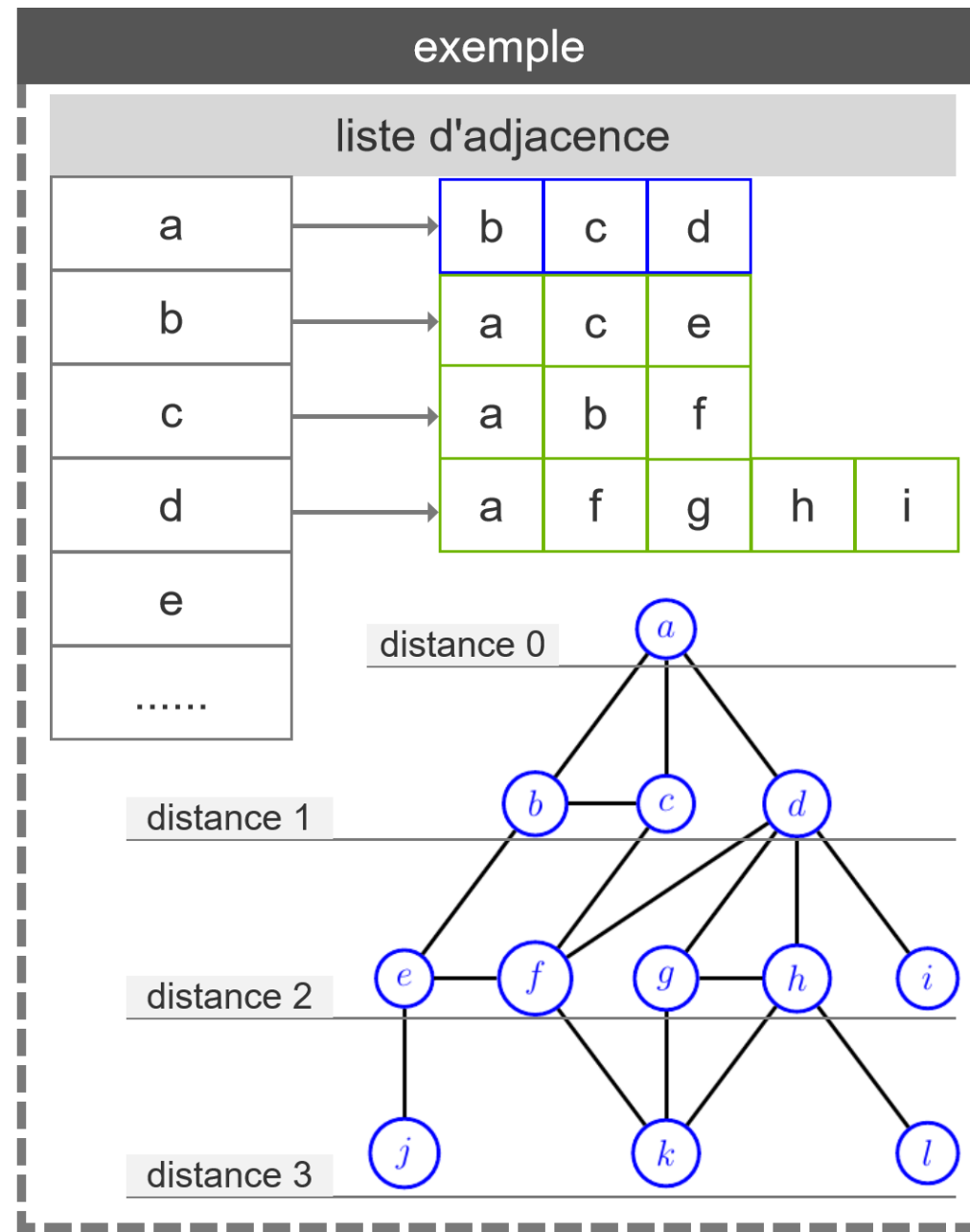
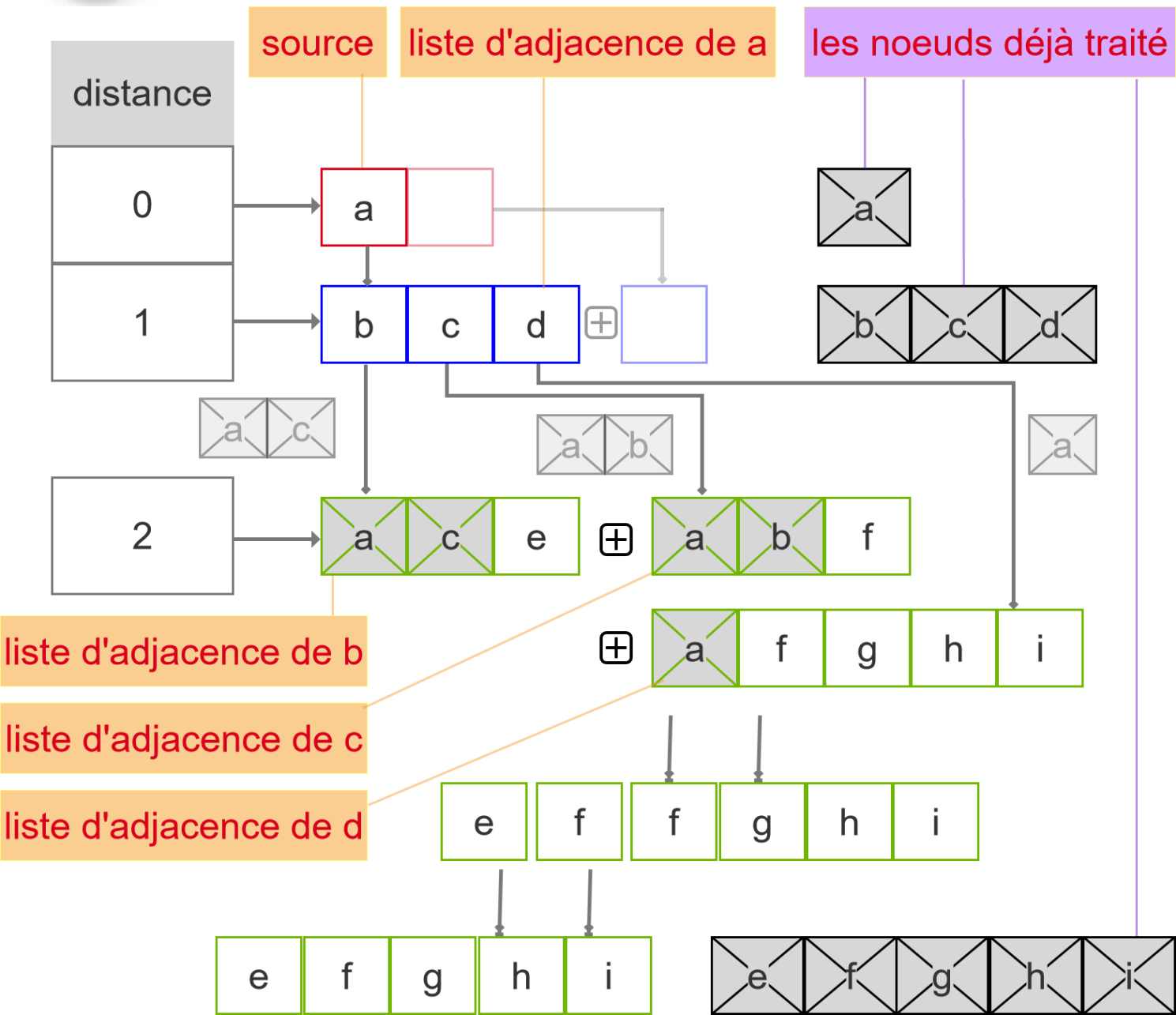


02

Partie algorithme

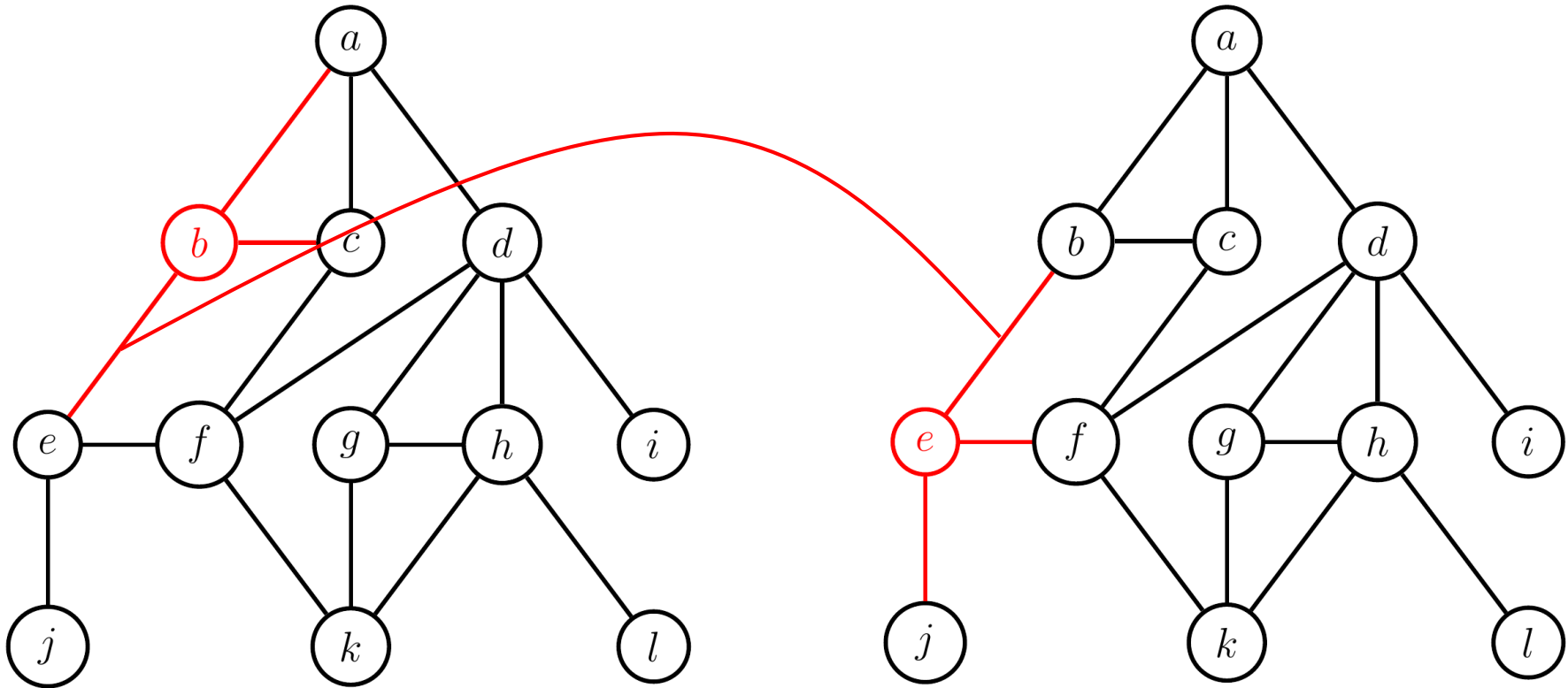


Distance



Complexité :

$\Theta(m)$ ($m = |E|$ = nombre d'arêtes)



Centralité de proximité

Algorithm 2: Calculer centralité de proximité

Entrée: Graphe : g ;

Sortie : Tableau : res ;

```
1 for noeud  $\in g$  do
2   somme  $\leftarrow 0.0$ ;
3    $d \leftarrow$  tableau de la liste de la distance pour noeud;
4   for  $n \in [1, \text{max distance de } d]$  do
5     | somme  $\leftarrow$  somme +  $n * \text{len}(d[n])$ ;
6   end
7   res[noeud]  $\leftarrow 1/\text{somme}$ ;
8 end
```

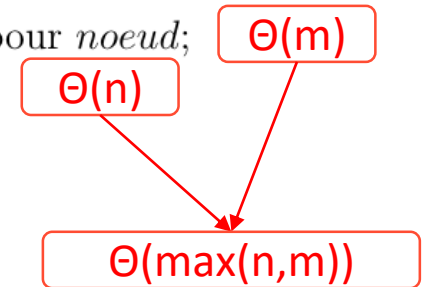
Centralité harmonique

Algorithm 2: Calculer centralité harmonique

Entrée: Graphe : g ;

Sortie : Tableau : res ;

```
1 for noeud  $\in g$  do n fois
2   somme  $\leftarrow 0.0$ ;
3    $d \leftarrow$  tableau de la liste de la distance pour noeud;
4   for  $n \in [1, \text{max distance de } d]$  do
5     | somme  $\leftarrow$  somme +  $\text{len}(d[n])/n$ ;
6   end
7   res[noeud]  $\leftarrow$  somme;
8 end
```



$$C_C(x) = \frac{1}{\sum_{y \in V, y \neq x} d(x, y)}$$

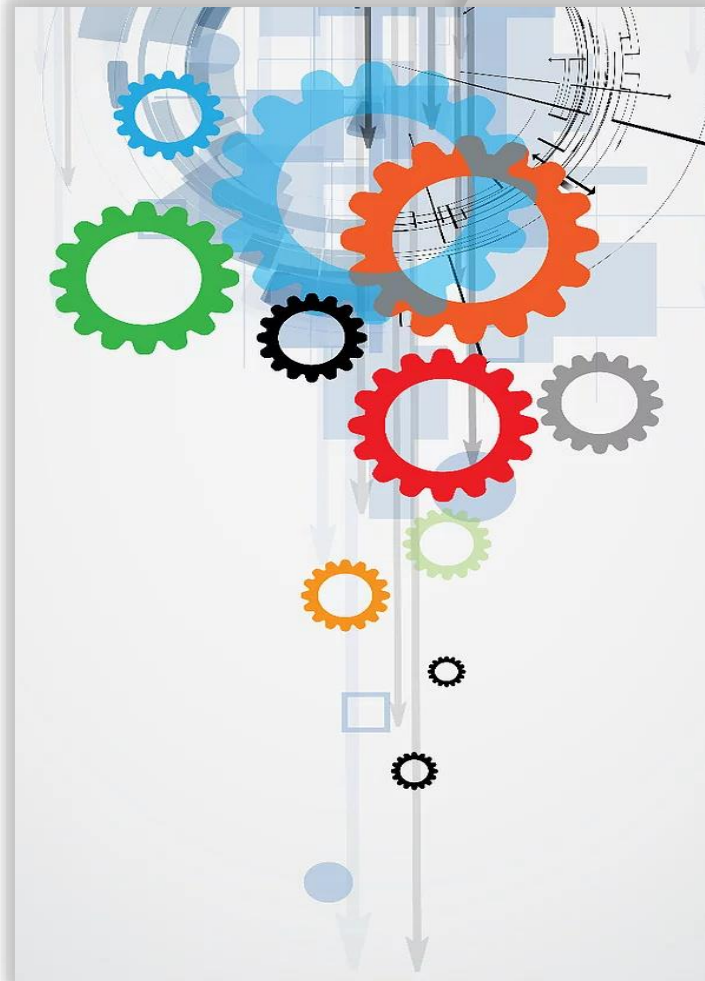
$$C_H(x) = \sum_{y \neq x} \frac{1}{d(y, x)}$$

Complexité : $\Theta (n * (\max(n,m)))$

($m = |E|$ = nombre d'arêtes ; $n = |V|$ = nombre de nœuds)

03

Partie implémentation



Python

Tableau --> Dictionnaire

Listes --> Listes
Python

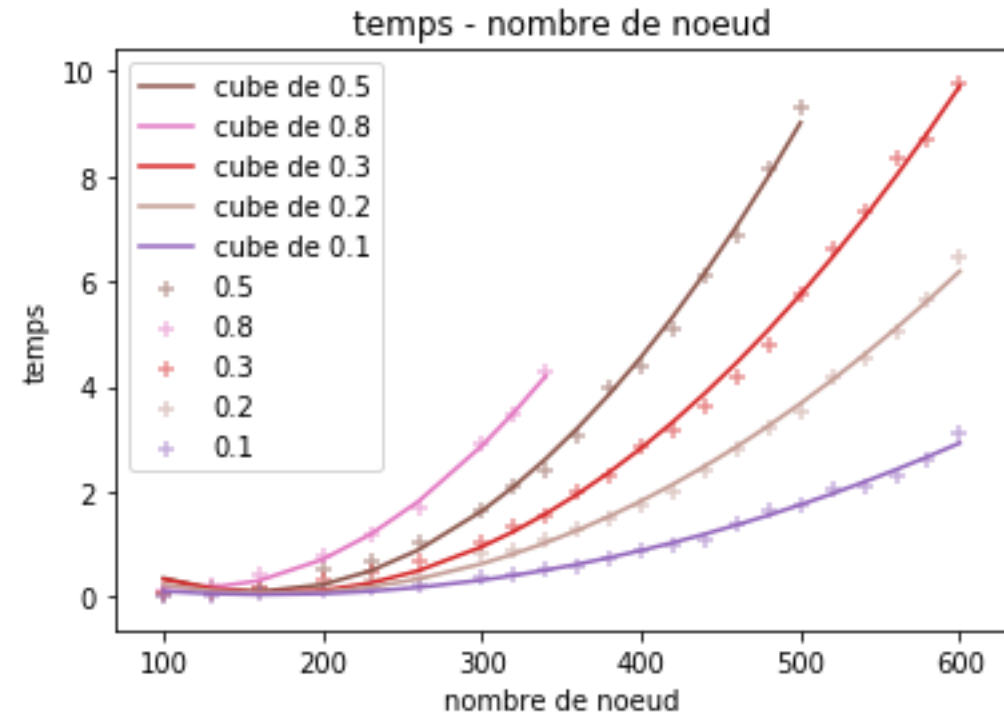
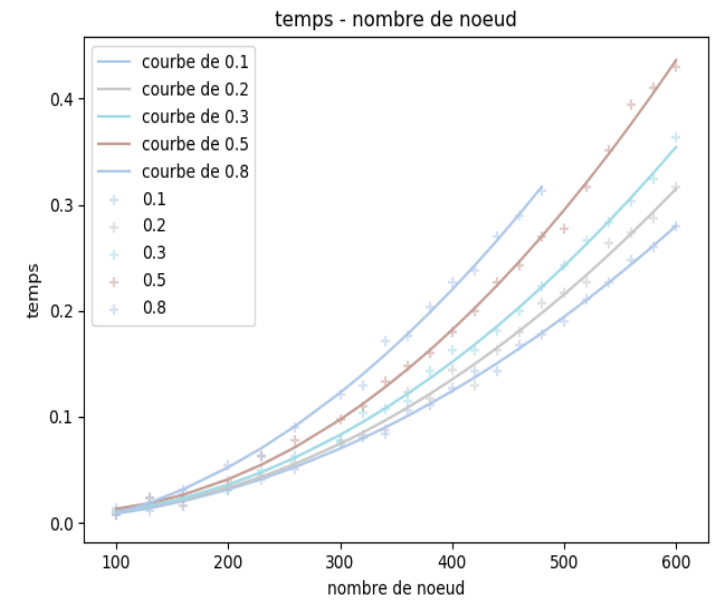
Graphe --> Dictionnaire
de listes Python

Graphe d'Erdos-Rényi

- n et m fixé
- les arrêtes choisis aléatoirement
- Souvent noté $G(n, M)$, avec M nombre d'arrêtes choisis

Expérience :

Densité ($2M/n(n-1)$) fixé.
 $\Theta(m) = \Theta(n^2)$



04

Partie analyse de données



Partie analyse de données

Motivation

Trouver quelles centralités sont corrélées et de conclure les spécialités de corrélation.

Explication pour Pearson et Spearman

- Si $-1 < r < 0$: forte relation linéaire négative entre x et y
- Si $r = 0$: absence de relation linéaire entre x et y
- Si $0 < r < 1$: forte relation linéaire positive entre x et y

Quelles centralités sont plus corrélées?



Coefficient Pearson

$$r_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$



Coefficient Spearman

Comparaison pratique des classements sur un jeux de données

Coefficients de Pearson et Spearman

Calculer et comparer

le graphe des métros de chicago

Il contient 584 noeuds et
638 arêtes

Ch et Cd

Centralité harmonique et
Centralité de degré

Ch et Cc

Centralité harmonique et
Centralité de proximité

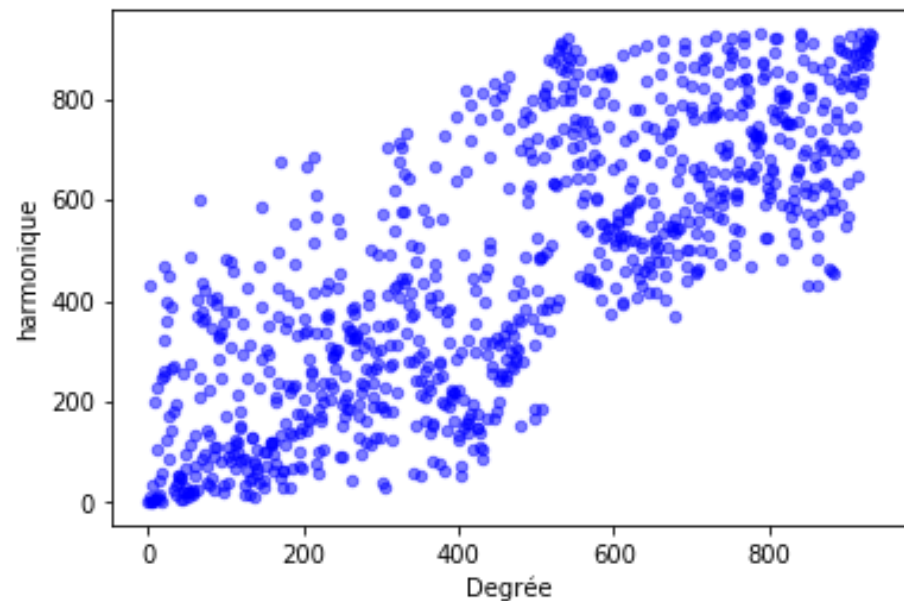
Cd et Cc

Centralité de degré et
Centralité de proximité

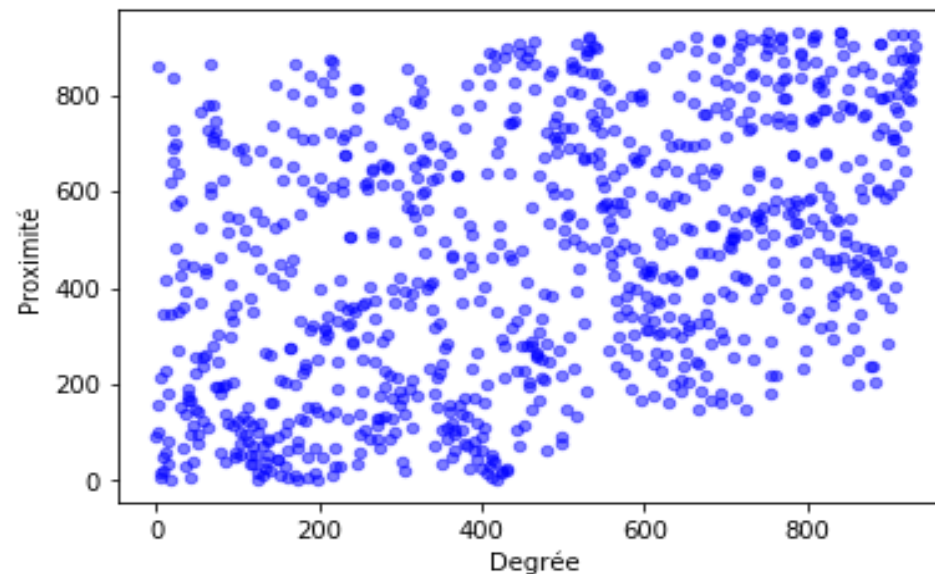
Résultat

Comparaison pratique des classements sur un jeu de données

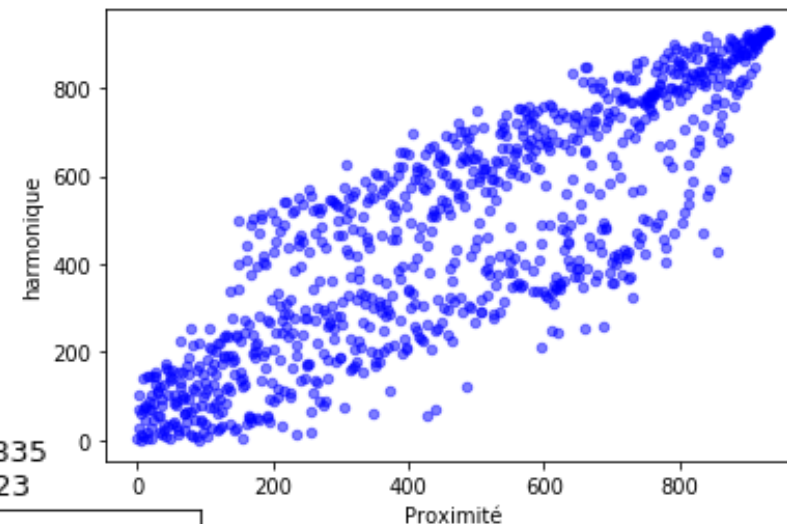
Spearman = 0.779930367476028
Pearson = 0.768500553213671



Spearman = 0.4288118272040835
Pearson = 0.3924431693667923



Spearman = 0.8462477091978801
Pearson = 0.8552061467288214

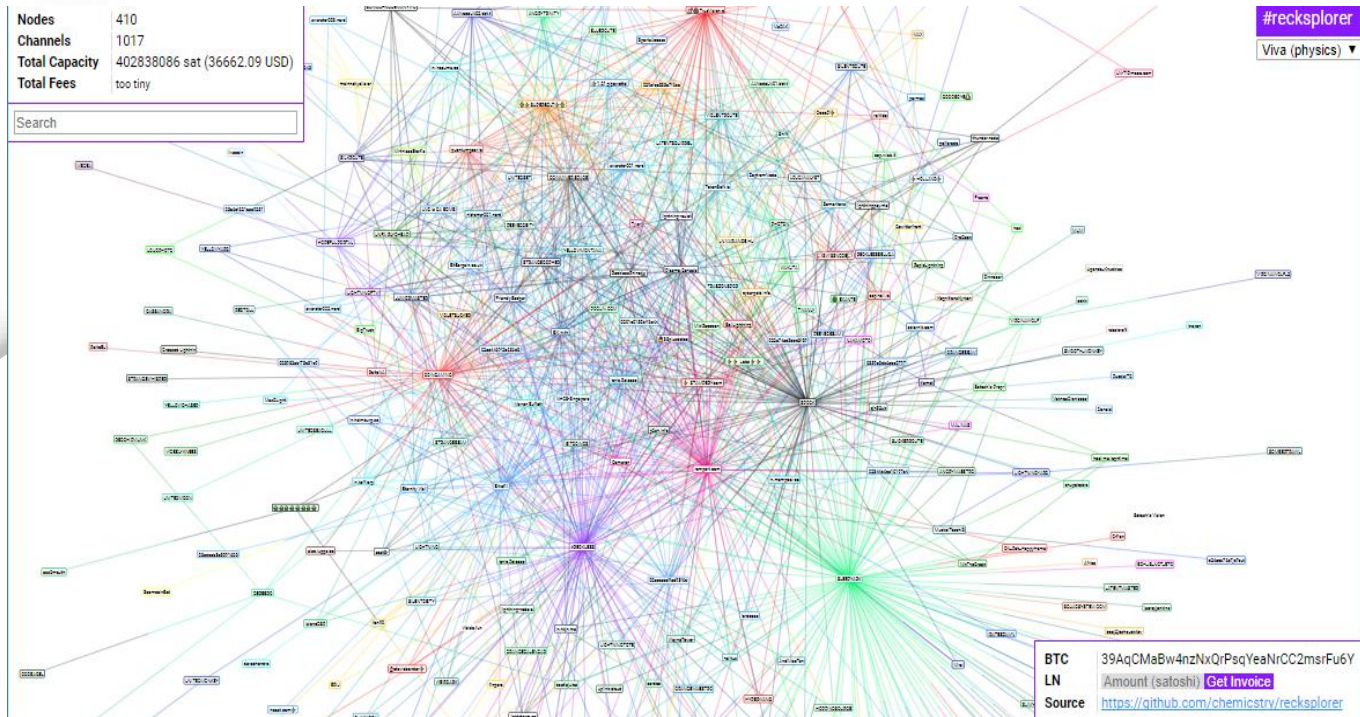


05

Résumé

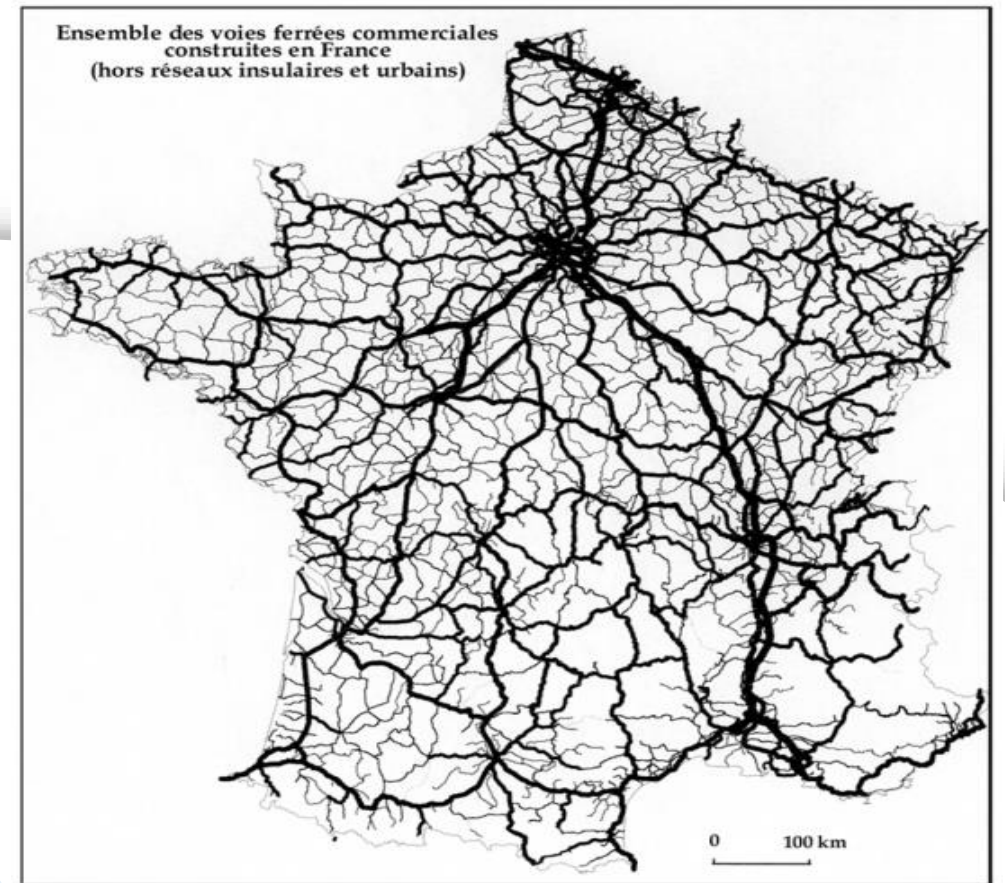


Caractérisation de centralité

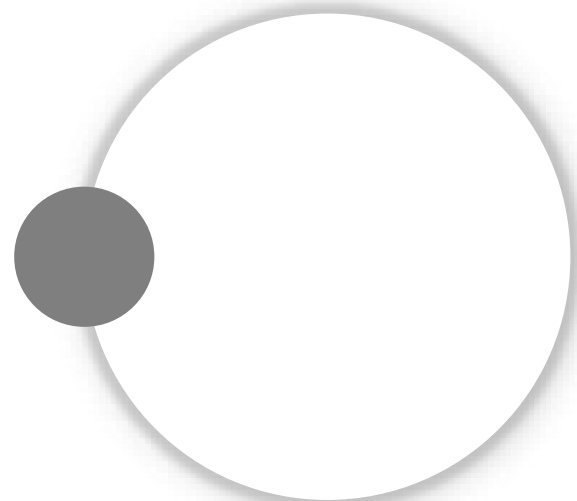


Caractérisation par un
flux de réseau

Caractérisation par la
structure de chemins



Notre point de vue



Merci beaucoup !

