Limites des centralités de degré, de proximité, d'intermédiarité

- On peut avoir une **centralité de degré** élevée mais on ne tient pas compte de la qualité des liens, ni de l'importance des voisins.

- On peut avoir une **centralité de proximité** artificiellement élevée pour des nœuds isolés qui ne sont pas connectés au reste du réseau.

- On peut avoir une **centralité d'intermédiarité** faible car il peut exister de nombreux chemins redondants dans le réseau mais le nœud peut être toujours important pour la connectivité du réseau.

Centralité de vecteur propre (eigenvector)
Bonacich (1972)

- Idée :
  - Donner du poids aux nœuds qui ont un degré élevé,
  - qui sont adjacents à des sommets ayant un degré élevé,
  - qui sont eux mêmes adjacent à des sommets qui ont un degré élevé,
  - …
- C'est une centralité qui prend en compte l'importance de nos voisins, des voisins de nos voisins, …
- Centralité globale : tient compte de la structure globale du réseau

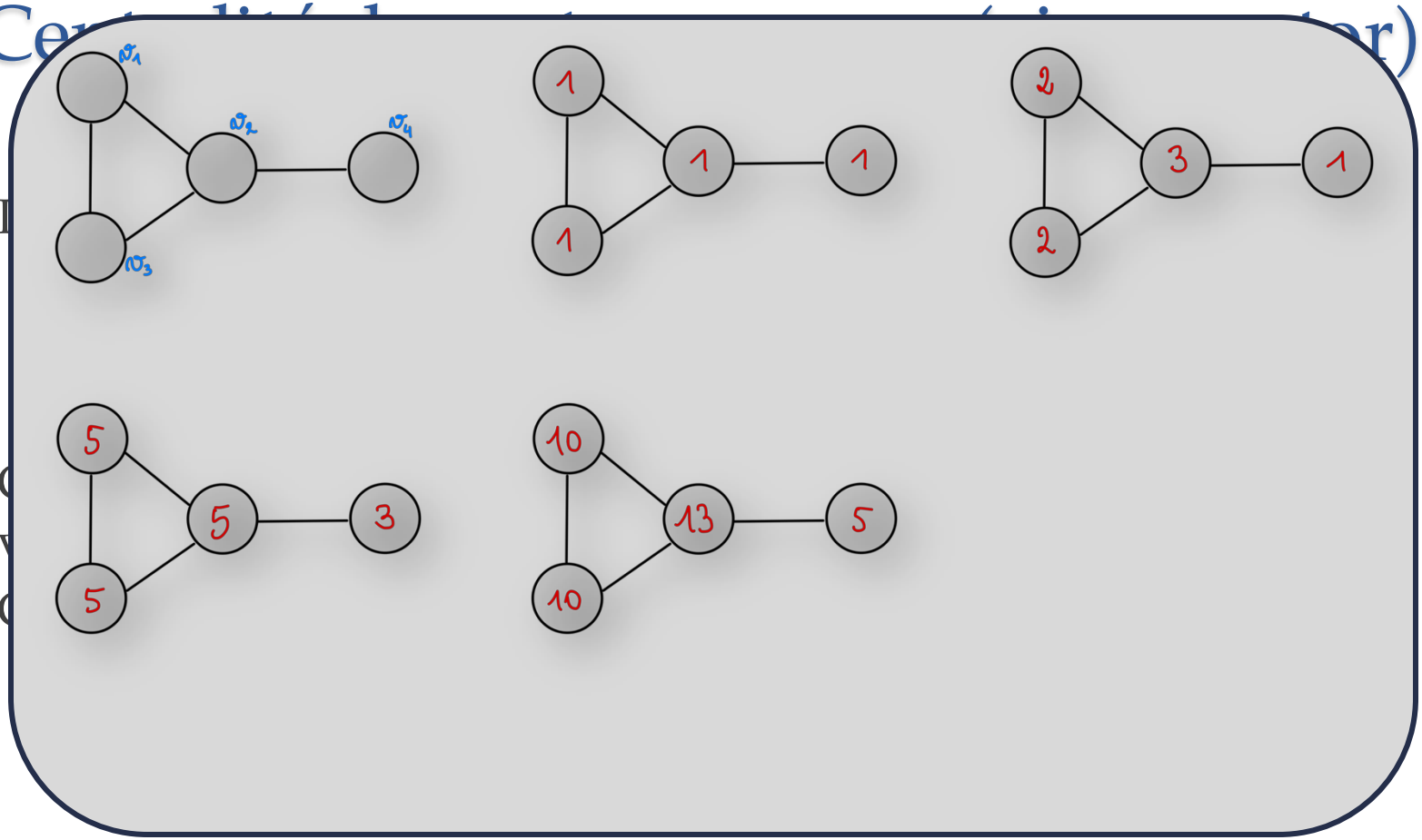
Centralité de vecteur propre (eigenvector)
Bonacich (1972)

- Idée :
  - Donner du poids aux nœuds qui ont un degré élevé,
  - qui sont adjacents à des sommets ayant un degré élevé,
  - qui sont eux mêmes adjacent à des sommets qui ont un degré élevé,
  - …
- C'est une centralité qui prend en compte l'importance de nos voisins, des voisins de nos voisins, …
- Centralité globale : tient compte de la structure globale du réseau

$$C_e(v_i) = \sum_{v_j \text{ voisin de } v_i} C_e(v_j)$$

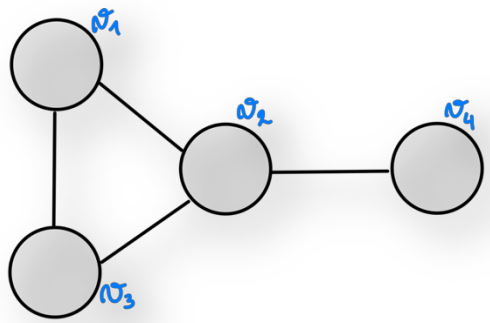
Centralité (vecteur propre)

- I
- C
  v
- C



$$C_e(v_i) = \sum_{v_j \text{ voisin de } v_i} C_e(v_j)$$

Centralité de vecteur propre (eigenvector)
Bonacich (1972)



$$\begin{cases} C_e(v_1) = C_e(v_2) + C_e(v_3) \\ C_e(v_2) = C_e(v_1) + C_e(v_3) + C_e(v_4) \\ C_e(v_3) = C_e(v_1) + C_e(v_2) \\ C_e(v_4) = C_e(v_2) \end{cases}$$

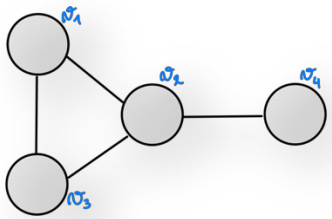


Solution unique et triviale :

$$\begin{cases} C_e(v_1) = 0 \\ C_e(v_2) = 0 \\ C_e(v_3) = 0 \\ C_e(v_4) = 0 \end{cases}$$

Centralité de vecteur propre (eigenvector)

Bonacich (1972)



Nouvelle idée

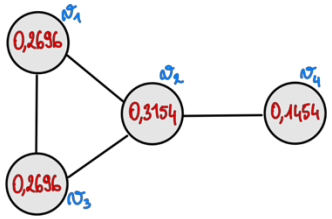
Résoudre le système suivant :

$$\begin{cases} \lambda\, C_e(v_1) = C_e(v_2) + C_e(v_3) \\ \lambda\, C_e(v_2) = C_e(v_1) + C_e(v_3) + C_e(v_4) \\ \lambda\, C_e(v_3) = C_e(v_1) + C_e(v_2) \\ \lambda\, C_e(v_4) = C_e(v_2) \end{cases}$$

- S'il existe un réel λ tel que le système possède une solution non nulle à ε près, on obtient une centralité pour chaque sommet qui dépend de la centralité de ses voisins à un facteur $1/\lambda$ près.

Centralité de vecteur propre (eigenvector)

Bonacich (1972)



Nouvelle idée

Résoudre le système suivant :

$$\begin{cases} \lambda\, C_e(v_1) = C_e(v_2) + C_e(v_3) \\ \lambda\, C_e(v_2) = C_e(v_1) + C_e(v_3) + C_e(v_4) \\ \lambda\, C_e(v_3) = C_e(v_1) + C_e(v_2) \\ \lambda\, C_e(v_4) = C_e(v_2) \end{cases}$$

Bonne nouvelle !

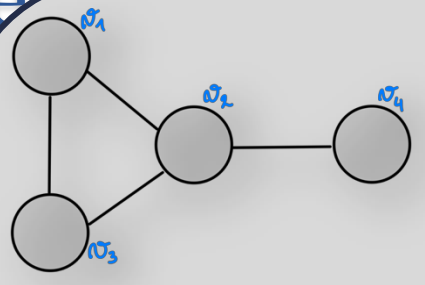
$$\lambda = 2.1701,\ C_e(v_1) = 0.2696,\ C_e(v_2) = 0.3154,\ C_e(v_3) = 0.2696,\ C_e(v_4) = 0.1454$$

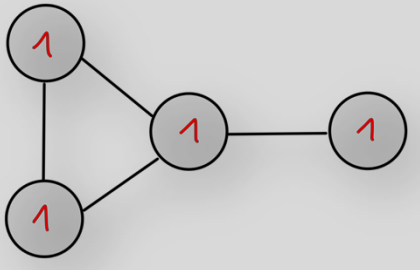
$$\begin{cases} 2.1701 \times 0.2696 = 0.3154 + 0.2696 \\ 2.1701 \times 0.3154 = 0.2696 + 0.2696 + 0.1454 \\ 2.1701 \times 0.2696 = 0.3154 + 0.2696 \\ 2.1701 \times 0.1454 = 0.3154 \end{cases}$$

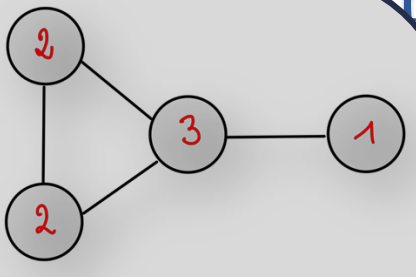
à 10^{-3} près

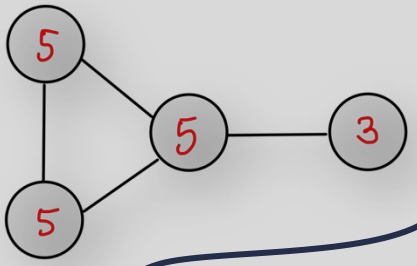


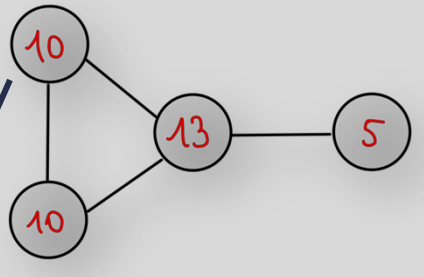
Centralité (vecteur propre)











$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$C_e(v_1) = a_{11}C_e(v_1) + a_{12}C_e(v_2) + a_{13}C_e(v_3) + a_{14}C_e(v_4)$$
$$= 0 \times 5 + 1 \times 5 + 1 \times 5 + 0 \times 3 = 10$$

$$C_e(v_i) = \sum_{v_j \text{ voisin de } v_i} C_e(v_j)$$

$$= \sum_{j=1}^{n} a_{ij}\, C_e(v_j)$$

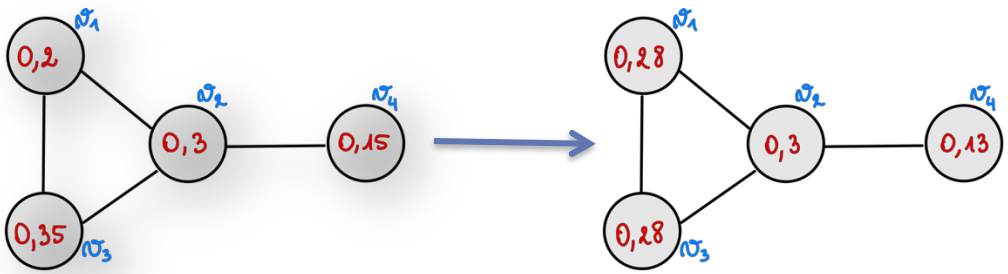
$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Centralité de vecteur propre (eigenvector)
Bonacich (1972)

- Facteur de normalisation :

$$C_e(v_i) = \sum_j a_{ij} C_e(v_j) \xrightarrow[\lambda = \sum_i \sum_j a_{ij} C_e(v_j)]{\text{Normalisation}} C_e(v_i) = \frac{1}{\lambda} \sum_j a_{ij} C_e(v_j)$$

- Notation matricielle : $C_e = \frac{1}{\lambda} A C_e$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad C_e = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.35 \\ 0.15 \end{pmatrix} \quad AC_e = \begin{pmatrix} 0.65 \\ 0.7 \\ 0.65 \\ 0.3 \end{pmatrix} \quad \frac{1}{\lambda} AC_e = \begin{pmatrix} 0.28 \\ 0.3 \\ 0.28 \\ 0.13 \end{pmatrix}$$

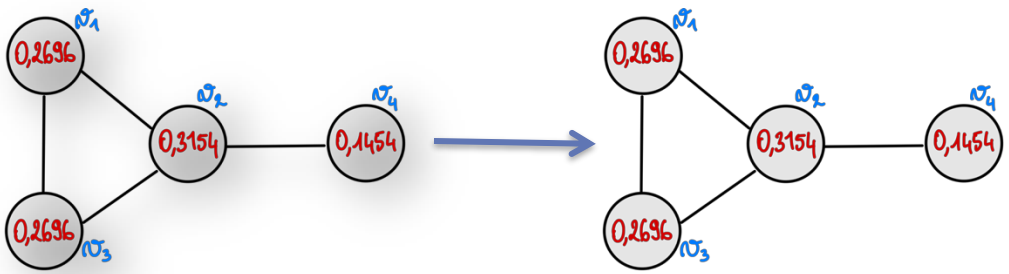
$$\Sigma \rightarrow \lambda = 2.3$$

Centralité de vecteur propre (eigenvector)
Bonacich (1972)

- Facteur de normalisation :

$$C_e(v_i) = \sum_j a_{ij} C_e(v_j) \xrightarrow[\lambda = \sum_i \sum_j a_{ij} C_e(v_j)]{\text{Normalisation}} C_e(v_i) = \frac{1}{\lambda} \sum_j a_{ij} C_e(v_j)$$

- Notation matricielle : $C_e = \frac{1}{\lambda} A C_e$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad C_e = \begin{pmatrix} 0.2696 \\ 0.3154 \\ 0.2696 \\ 0.1454 \end{pmatrix} \quad AC_e = \begin{pmatrix} 0.585 \\ 0,6846 \\ 0.585 \\ 0.3154 \end{pmatrix} \quad \frac{1}{2.1701} AC_e = \begin{pmatrix} 0.2696 \\ 0.3154 \\ 0.2696 \\ 0.1454 \end{pmatrix}$$

Centralité de vecteur propre (eigenvector)
Bonacich (1972)

- Facteur de normalisation :

$$C_e(v_i) = \sum_j a_{ij} C_e(v_j) \xrightarrow[\lambda = \sum_i \sum_j a_{ij} C_e(v_j)]{\text{Normalisation}} C_e(v_i) = \frac{1}{\lambda} \sum_j a_{ij} C_e(v_j)$$

- Notation matricielle : $C_e = \frac{1}{\lambda} A C_e$
- Algèbre matriciel :
  - Valeur propre de A : λ
  - Vecteur propre associé : X

$$\lambda X = AX$$

- Il existe n valeurs propres. Des théorèmes mathématiques assurent que la plus grande valeur propre donne un vecteur propre positif et unique.

- Méthode algorithmique pour déterminer une valeur approchée de la plus grande valeur proche et du vecteur propre associé :
  Puissance itérée (power iteration)

Centralité de vecteur propre (eigenvector)
Algorithme de puissance itérée

- Permet d'approximer la plus grande valeur propre d'une matrice et un vecteur propre associé.
- Algorithme simple, mais qui ne converge pas très vite, à utiliser sur des matrices creuses.

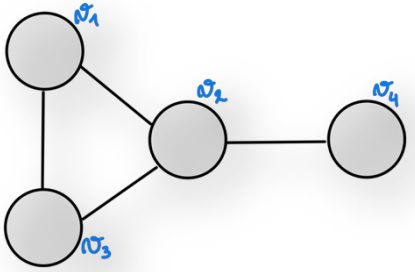
Algorithme de puissance itérée

1. Commencer par associer une centralité aléatoire C_e à tous les sommets telle que **toutes les valeurs somment à 1**.
2. Recalculer la centralité de chaque nœud comme la **somme des centralités de leurs voisins**, i.e. calculer AC_e et affecter le résultat dans C_e .
3. Normaliser le vecteur-colonne C_e **en divisant tous les éléments par la somme des valeurs**.
4. Recommencer les étapes 2 et 3 jusqu'à ce que les valeurs du vecteur-colonne ne varient plus (à ε près).

Centralité de vecteur propre (eigenvector)

Algorithme de puissance itérée

$$C_e = \frac{1}{\lambda} A C_e$$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C_e^0 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

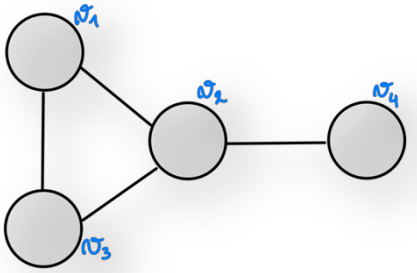
$$C_e^1 = AC_e^0 = \begin{pmatrix} 1/2 \\ 3/4 \\ 1/2 \\ 1/4 \end{pmatrix} \longrightarrow \overline{C_e^1} = \frac{1}{2} C_e^1 = \begin{pmatrix} 1/4 \\ 3/8 \\ 1/4 \\ 1/8 \end{pmatrix}$$

$$C_e^2 = A\overline{C_e^1} = \begin{pmatrix} 5/8 \\ 5/8 \\ 5/8 \\ 3/8 \end{pmatrix} \longrightarrow \overline{C_e^2} = \frac{4}{9} C_e^2 = \begin{pmatrix} 5/18 \\ 5/18 \\ 5/18 \\ 1/6 \end{pmatrix}$$

Centralité de vecteur propre (eigenvector)

Algorithme de puissance itérée

$$C_e = \frac{1}{\lambda} A C_e$$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C_e^0 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

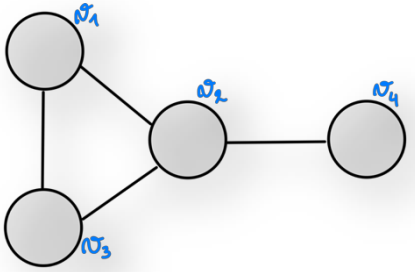
$$C_e^2 = A\overline{C_e^1} = \begin{pmatrix} 5/8 \\ 5/8 \\ 5/8 \\ 3/8 \end{pmatrix} \longrightarrow \overline{C_e^2} = \frac{4}{9} C_e^2 = \begin{pmatrix} 5/18 \\ 5/18 \\ 5/18 \\ 1/6 \end{pmatrix}$$

$$C_e^3 = A\overline{C_e^2} = \begin{pmatrix} 5/9 \\ 13/18 \\ 5/9 \\ 5/18 \end{pmatrix} \longrightarrow \overline{C_e^3} = \frac{9}{19} C_e^3 = \begin{pmatrix} 5/19 \\ 13/38 \\ 5/19 \\ 5/38 \end{pmatrix}$$

Centralité de vecteur propre (eigenvector)

Algorithme de puissance itérée

$$C_e = \frac{1}{\lambda} A C_e$$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C_e^0 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

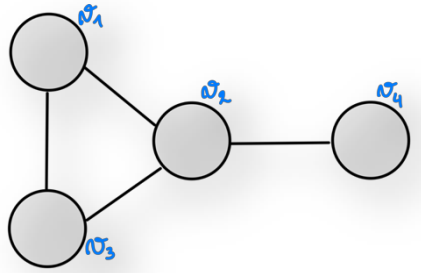
$$C_e^3 = A\overline{C_e^2} = \begin{pmatrix} 5/9 \\ 13/18 \\ 5/9 \\ 5/18 \end{pmatrix} \longrightarrow \overline{C_e^3} = \frac{9}{19} C_e^3 = \begin{pmatrix} 5/19 \\ 13/38 \\ 5/19 \\ 5/38 \end{pmatrix}$$

$$C_e^4 = A\overline{C_e^3} = \begin{pmatrix} 23/38 \\ 25/38 \\ 23/38 \\ 13/38 \end{pmatrix} \longrightarrow \overline{C_e^4} = \frac{19}{42} C_e^4 = \begin{pmatrix} 23/84 \\ 25/84 \\ 23/84 \\ 13/84 \end{pmatrix}$$

Centralité de vecteur propre (eigenvector)

Algorithme de puissance itérée

$$C_e = \frac{1}{\lambda} A C_e$$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C_e^0 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

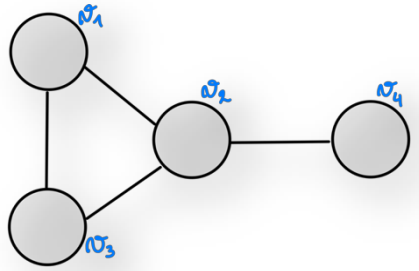
$$C_e^{20} = A\overline{C_e^{19}} = \begin{pmatrix} 0.585 \\ 0.6846 \\ 0,585 \\ 0,3154 \end{pmatrix} \rightarrow \overline{C_e^{20}} = \frac{1}{2.17} C_e^{20} = \begin{pmatrix} 0.2696 \\ 0.3155 \\ 0.2696 \\ 0.1453 \end{pmatrix}$$

Valeurs égales à 10^{-2} près

$$C_e^{21} = A\overline{C_e^{20}} = \begin{pmatrix} 0.5851 \\ 0,6845 \\ 0.5851 \\ 0.3155 \end{pmatrix} \rightarrow \overline{C_e^{21}} = \frac{1}{2.1701} C_e^{21} = \begin{pmatrix} 0.2696 \\ 0.3154 \\ 0.2696 \\ 0.1454 \end{pmatrix}$$

Que calcule la centralité de vecteur propre ?

$$C_e = \frac{1}{\lambda} A C_e$$



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C_e^0 = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

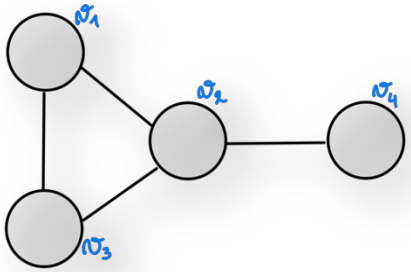
$$\frac{1}{\lambda_1} A C_e \longrightarrow \frac{1}{\lambda_2} A \times (\frac{1}{\lambda_1} A C_e) = \frac{1}{\lambda_1 \lambda_2} A^2 C_e \longrightarrow \frac{1}{\lambda_3} A \times (\frac{1}{\lambda_1 \lambda_2} A^2 C_e) = \frac{1}{\lambda_1 \lambda_2 \lambda_3} A^3 C_e$$

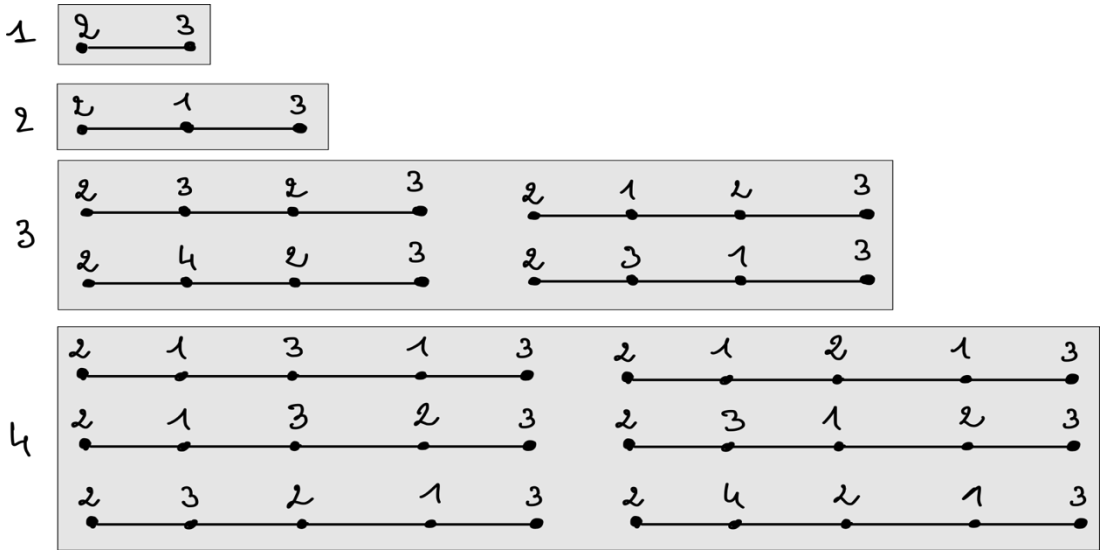
$$\longrightarrow \frac{1}{\lambda_4} A \times (\frac{1}{\lambda_1 \lambda_2 \lambda_3} A^3 C_e) = \frac{1}{\lambda_1 \lambda_2 \lambda_3 \lambda_4} A^4 C_e \longrightarrow \cdots \longrightarrow \frac{1}{\prod_{i=1}^{q} \lambda_i} A^q C_e$$

À quoi correspond la puissance $q^{\text{ième}}$ de la matrice d'adjacence ?

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \quad A^3 = \begin{pmatrix} 2 & 4 & 3 & 1 \\ 4 & 2 & 4 & 3 \\ 3 & 4 & 2 & 1 \\ 1 & 3 & 1 & 0 \end{pmatrix} \quad A^4 = \begin{pmatrix} 7 & 6 & 6 & 4 \\ 6 & 11 & 6 & 2 \\ 6 & 6 & 7 & 4 \\ 4 & 2 & 4 & 3 \end{pmatrix}$$

Ex. avec la valeur en ligne 2, colonne 3 de A^q :
nombre de chemins de longueur q reliant le sommet v_2 au sommet v_3

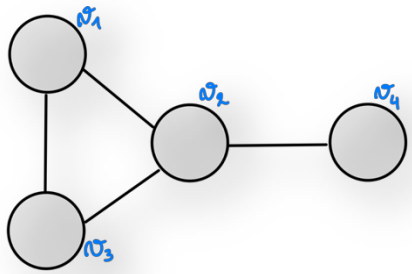




À quoi correspond la puissance $q^{\text{ième}}$ de la matrice d'adjacence ?

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \quad A^3 = \begin{pmatrix} 2 & 4 & 3 & 1 \\ 4 & 2 & 4 & 3 \\ 3 & 4 & 2 & 1 \\ 1 & 3 & 1 & 0 \end{pmatrix} \quad A^4 = \begin{pmatrix} 7 & 6 & 6 & 4 \\ 6 & 11 & 6 & 2 \\ 6 & 6 & 7 & 4 \\ 4 & 2 & 4 & 3 \end{pmatrix}$$

Ex. avec la valeur en ligne 2, colonne 3 de A^q :
nombre de chemins de longueur q reliant le sommet v_2 au sommet v_3



Et $A^q C_e$?

$$A^4 C_e = \begin{pmatrix} 7 & 6 & 6 & 4 \\ 6 & 11 & 6 & 2 \\ 6 & 6 & 7 & 4 \\ 4 & 2 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} 23/4 \\ 25/4 \\ 23/4 \\ 13/4 \end{pmatrix}$$

Nombre de chemins

Sachant que $C_e = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$ à l'initialisation, $A^q C_e$ calcule le nombre moyen de chemins de longueur q qui atteingnent chacun des sommets depuis chaque sommet v.

$$A^q C_e = \begin{pmatrix} \#\text{ moyen de chemins de longueur } q \text{ qui atteignent } v_1 \text{ depuis chacun des sommets} \\ \#\text{ moyen de chemins de longueur } q \text{ qui atteignent } v_2 \text{ depuis chacun des sommets} \\ \#\text{ moyen de chemins de longueur } q \text{ qui atteignent } v_3 \text{ depuis chacun des sommets} \\ \#\text{ moyen de chemins de longueur } q \text{ qui atteignent } v_4 \text{ depuis chacun des sommets} \end{pmatrix}$$

Nous avons :

$$A^{21} = \begin{pmatrix} 3180604 & 3722857 & 3180605 & 1714550 \\ 3722857 & 4352902 & 3722857 & 2008307 \\ 3180605 & 3722857 & 3180604 & 1714550 \\ 1714550 & 2008307 & 1714550 & 923802 \end{pmatrix}$$

Et donc :

$$A^{21} C_e = \begin{pmatrix} 2949654 \\ 3451730.75 \\ 2949654 \\ 1590302.25 \end{pmatrix}$$

Nombre de chemins

$\sum = 10941341$

Nous avons :

$$A^{21} = \begin{pmatrix} 3180604 & 3722857 & 3180605 & 1714550 \\ 3722857 & 4352902 & 3722857 & 2008307 \\ 3180605 & 3722857 & 3180604 & 1714550 \\ 1714550 & 2008307 & 1714550 & 923802 \end{pmatrix}$$

Et donc :

$$A^{21}C_e = \begin{pmatrix} 2949654 \\ 3451730.75 \\ 2949654 \\ 1590302.25 \end{pmatrix}$$

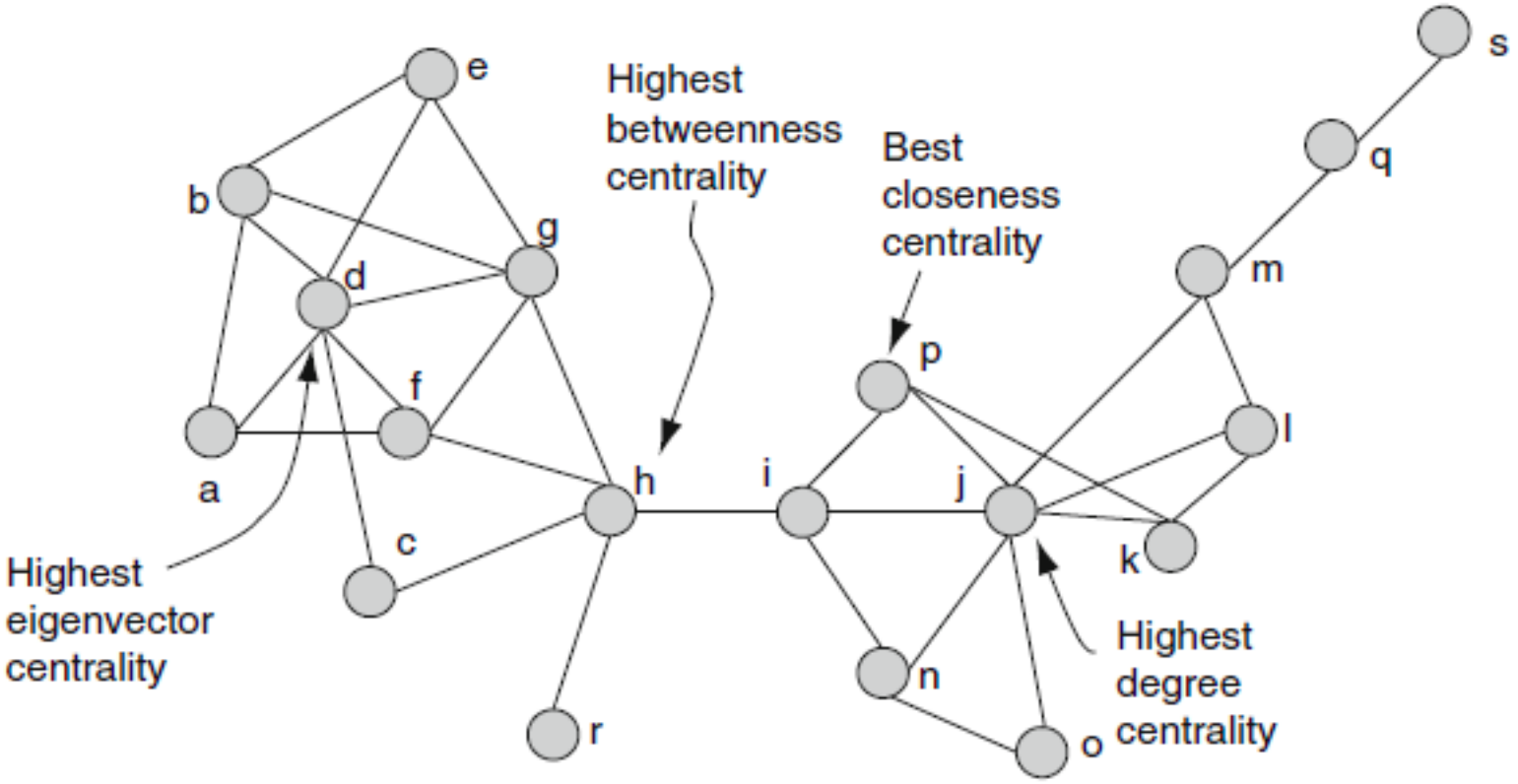
Après normalisation :

$$\frac{1}{10941341} A^{21}C_e = \begin{pmatrix} 0.2696 \\ 0.3154 \\ 0.2696 \\ 0.1454 \end{pmatrix}$$

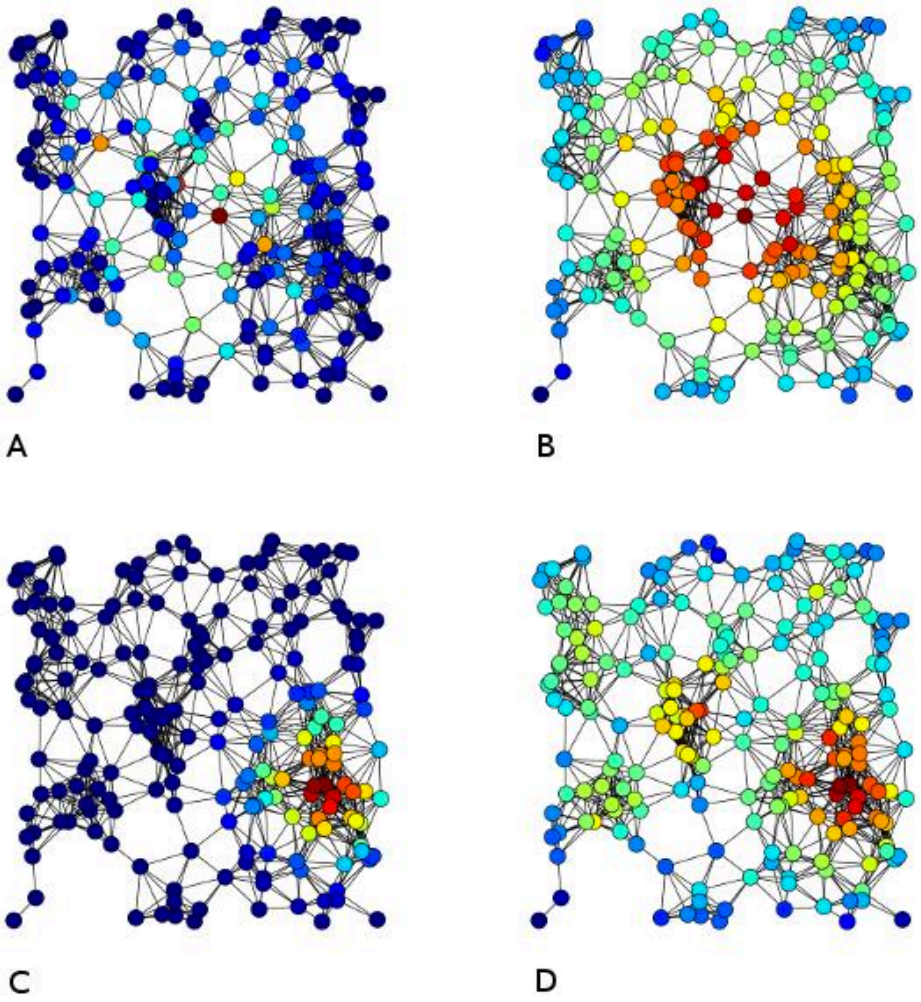
Centralité de vecteur propre : tient compte de l'importance des voisins, des voisins des voisins, …

Un sommet est important si beaucoup de chemins arrivent jusqu'à lui, ce qui donne de l'importance à ses voisins qui sont atteignables à un pas de plus.

Comparaison des paramètres



Comparaison des paramètres



A. Intermédiarité

B. Proximité

C. Vecteur propre

D. Degré