



ECOLE CENTRALE CASABLANCA

## THÉORIE DES GRAPHS

---

# Projet final : réseaux sociaux et graphes

---

**Membres :**

EL MRABET Aimane  
EL BARHICHI Mohammed

**Sous la supervision de :**

M. Abdessamad Ait Elcadi

15 décembre 2023

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Description du problème</b>	<b>2</b>
<b>3</b>	<b>Premier réseau : réseau 0</b>	<b>2</b>
3.1	Centralité en terme d'importance du voisinage . . . . .	2
3.1.1	Calcul du PageRank pour $\alpha = 1$ et $\beta = 0$ . . . . .	3
3.1.2	Calcul du PageRank pour $\alpha = 0,85$ et $\beta = 1$ . . . . .	3
3.1.3	Calcul du PageRank pour $\alpha = 0$ et $\beta = 1$ . . . . .	4
3.1.4	Discussion des effets des différentes valeurs de $\alpha$ et de $\beta$ pour ce réseau particulier. . . . .	4
3.1.5	Quelle valeur de $\alpha$ choisir pour calculer les valeurs du PageRank ?	5
3.2	Centralité d'intermédiarité . . . . .	5
3.3	Centralité de proximité . . . . .	6
3.4	Table de synthèse . . . . .	6
<b>4</b>	<b>Second réseau : réseau des membres d'un club de Karaté</b>	<b>7</b>
<b>5</b>	<b>Troisième réseau : exemple d'un réseau de type Facebook</b>	<b>8</b>
<b>6</b>	<b>Quatrième réseau : Réseau d'un groupe de chercheurs-physiciens</b>	<b>9</b>
<b>7</b>	<b>Étude qualitative des résultats obtenus et comparaison des réseaux</b>	<b>11</b>

# 1 Introduction

Dans l'ère numérique contemporaine, les réseaux sociaux occupent une place centrale dans notre quotidien, agissant comme des vecteurs essentiels de communication, de partage d'informations et de création de liens sociaux. L'étude de la dynamique au sein de ces réseaux constitue une discipline cruciale pour comprendre les schémas d'interaction et les influences qui émergent au sein de ces vastes ensembles interconnectés.

Au cœur de cette analyse réside le concept fondamental de **"centralité"**. La centralité dans un réseau offre une perspective clé pour quantifier et évaluer l'importance relative des différents acteurs au sein d'un système complexe. Elle permet de discerner les individus, les nœuds, ou les entités qui jouent un rôle pivot, exercent une influence significative, ou occupent une position stratégique au sein du réseau social.

Tout au long de ce rapport, nous essayerons de comprendre en profondeur cette notion de centralité et ses différents types, à travers l'étude de plusieurs cas de figures.

## 2 Description du problème

Ce projet vise à étudier et comparer trois types de centralité : La centralité en termes d'importance du voisinage (Edge Centrality), la centralité d'intermédiarité (Betweenness Centrality) et enfin la centralité de proximité (Closeness Centrality). Et cela à travers l'étude de quatre exemples de réseaux différents : un réseau simple noté réseau 0, un réseau de type Facebook, un réseau des membres d'un club de karaté et enfin n réseau d'un groupe de chercheurs physiciens.

## 3 Premier réseau : réseau 0

### 3.1 Centralité en terme d'importance du voisinage

Dans cette partie, nous étudierons le réseau 0 ci-dessous et nous nous intéresserons à la centralité en terme de l'importance du voisinage, et plus particulièrement le Page-Rank.

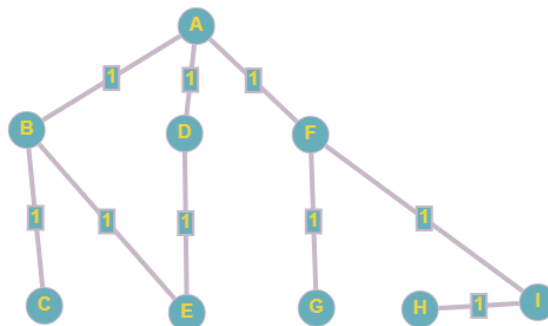


FIGURE 1 – Graphe du réseau 0

Nous procédons par la suite au calcul des valeurs du Pagerank pour différentes valeurs de  $\alpha$  et de  $\beta$ .

### 3.1.1 Calcul du Pagerank pour $\alpha = 1$ et $\beta = 0$

Afin de calculer les valeurs du Pagerank pour  $\alpha = 1$  et  $\beta = 0$ , nous utilisons la méthode des puissances vue en cours.

Nous implémentons cette méthode en utilisant un code python (voir code dans le fichier .ipynb) et nous remarquons que l'algorithme de la méthode des puissances ne converge pas et qu'après un grand nombre d'itérations, nous avons deux séries de valeurs qui alternent entre elles.

Ces deux séries de valeurs du Pagerank sont :

Noeud	Valeur du Pagerank	Rank
A	0.185185	1
B	0.148148	2
C	0.061728	7
D	0.098765	6
E	0.123457	4
F	0.148148	2
G	0.061728	7
H	0.049383	8
I	0.123457	4

TABLE 1 – Première série des valeurs du Pagerank

Noeud	Valeur du Pagerank	Rank
A	0.148148	3
B	0.185185	1
C	0.049383	7
D	0.123457	8
E	0.098765	4
F	0.185185	1
G	0.049383	7
H	0.061728	6
I	0.098766	4

TABLE 2 – Deuxième série des valeurs du Pagerank

### 3.1.2 Calcul du Pagerank pour $\alpha = 0,85$ et $\beta = 1$

Dans ce cas de figure, nous utilisons la formule algébrique qui donne la valeur du Pagerank et qui est :  $C_p = \beta \left( \mathbf{I} - \alpha A^T D^{-1} \right)^{-1} \cdot \mathbf{1}$ .

Nous avons également implémenté cette formule dans python (voir code dans le fichier .ipynb) et nous obtenons les valeurs du PageRank suivantes :

Noeud	Valeur du PageRank	Rank
A	9.15056407483829	3
B	9.397704812491918	2
C	3.6626830302060434	8
D	6.284424236816782	6
E	6.333563330853176	5
F	9.942355567006281	1
G	3.817000743985113	7
H	4.105245113414751	6
I	7.306459090387651	4

TABLE 3 – Les valeurs du PageRank pour  $\alpha = 0.85$  et  $\beta = 1$

### 3.1.3 Calcul du PageRank pour $\alpha = 0$ et $\beta = 1$

Dans ce cas de figure, nous utilisons la formule algébrique qui donne la valeur du PageRank et qui est :  $\mathbf{C}_p = \beta (\mathbf{I} - \alpha A^T D^{-1})^{-1} \cdot \mathbf{1}$  afin de calculer les valeurs du PageRank pour  $\alpha = 0$  et  $\beta = 1$ .

Nous avons utilisé le même code que la section précédente (voir code dans le fichier .ipynb) et nous obtenons les valeurs du PageRank suivantes :

Noeud	Valeur du PageRank	Rank
A	1	1
B	1	1
C	1	1
D	1	1
E	1	1
F	1	1
G	1	1
H	1	1
I	1	1

TABLE 4 – Les valeurs du PageRank pour  $\alpha = 0$  et  $\beta = 1$

### 3.1.4 Discussion des effets des différentes valeurs de $\alpha$ et de $\beta$ pour ce réseau particulier.

Pour  $\alpha = 1$  et  $\beta = 0$ , l'algorithme de la méthode des puissances ne converge pas pour le réseau 0. Cela signifie qu'il n'y a pas de valeur unique de PageRank stable pour les nœuds. Les deux séries de valeurs du PageRank qui alternent indiquent une instabilité dans l'estimation de l'importance des nœuds. Cela peut être dû à des problèmes

de convergence de l'algorithme.

En revanche, pour  $\alpha = 0,85$  et  $\beta = 1$ , la formule algébrique du PageRank est utilisée pour calculer les valeurs du PageRank. Les valeurs obtenues reflètent l'importance relative des nœuds dans le réseau 0. Les nœuds ayant les valeurs de PageRank les plus élevées sont considérés comme les plus importants en termes de centralité. Cela peut indiquer qu'ils ont de nombreux liens sortants vers d'autres nœuds importants dans le réseau.

Lorsque  $\alpha = 0$  et  $\beta = 1$ , toutes les valeurs du PageRank sont égales à 1 pour le réseau 0. Cela signifie que tous les nœuds ont la même importance en termes de centralité. Dans ce cas, le PageRank ne tient pas compte de la structure du graphe ni des liens sortants des nœuds. Tous les nœuds sont considérés comme également importants, ce qui peut ne pas refléter la réalité du réseau.

### 3.1.5 Quelle valeur de $\alpha$ choisir pour calculer les valeurs du PageRank ?

Afin de garantir la convergence de l'algorithme du PageRank,  $\alpha$  doit être choisi de telle façon qu'il soit inférieur à l'inverse de la plus grande valeur propre de la matrice d'adjacence.

En d'autres termes, notons  $A$  la matrice d'adjacence du graphe  $G$  et  $\lambda_{max}$  la valeur propre la plus grande de  $A$ . Alors, la valeur de  $\alpha$  choisie doit satisfaire :

$$\alpha < \frac{1}{\lambda_{max}}$$

## 3.2 Centralité d'intermédierité

Après l'implémentation des différents codes qui aboutissent au calcul de la centralité d'intermédierité (voir le fichier .ipynb), nous obtenons, pour le réseau 0, les résultats suivants :

Noeud	Centralité d'intermediarité	Rank
A	0.5714285714285715	2
B	0.33035714285714285	3
C	0.0	7
D	0.08928571428571426	5
E	0.03571428571428571	6
F	0.5803571428571429	1
G	0.0	7
H	0.0	7
I	0.2410714285714285	4

TABLE 5 – Les valeurs de la centralité d'intermédierité du réseau 0

### 3.3 Centralité de proximité

Après l'implémentation des différents codes qui aboutissent au calcul de la centralité de proximité (voir le fichier .ipynb), nous obtenons, pour le réseau 0, les résultats suivants :

Noeud	Centralité de proximité	Rank
A	0.6428571428571428	1
B	0.5294117647058824	3
C	0.375	8
D	0.47368421052631576	4
E	0.40909090909090906	6
F	0.6	2
G	0.40909090909090906	6
H	0.3333333333333333	9
I	0.44999999999999996	5

TABLE 6 – Les valeurs de la centralité de proximité du réseau 0

### 3.4 Table de synthèse

Afin de résumer cette partie qui porte sur l'étude de différents types de centralité des noeuds du réseau 0, nous élaborons la table de synthèse de ce réseau. Nous signalons que pour les valeurs du Pagerank, nous considérons  $\alpha = 0.85$  et  $\beta = 1$ .

Caractéristiques	Valeur	Noeud concerné
Nombre de noeuds	9	N/A
Nombre d'arêtes	9	N/A
Degré - min	1	C, G, H
Degré - max	3	A, B, F
Degré - moyen	2	N/A
Centralité (PageRank) - min	3.66	C
Centralité (PageRank) - max	9.94	F
Centralité (PageRank) - moyenne	6.66	N/A
Centralité d'intermédarité - min	0.0	C, G, H
Centralité d'intermédarité - max	0.58	F
Centralité d'intermédarité - moyenne	0.205	N/A
Centralité de proximité - min	0.333	H
Centralité de proximité - max	0.64	A
Centralité de proximité - moyenne	0.46	N/A

TABLE 7 – Table de synthèse du réseau 0

## 4 Second réseau : réseau des membres d'un club de Karaté

Nous commençons cette partie par visualiser, à l'aide de python, le graphe du réseau des membres du club de karaté.

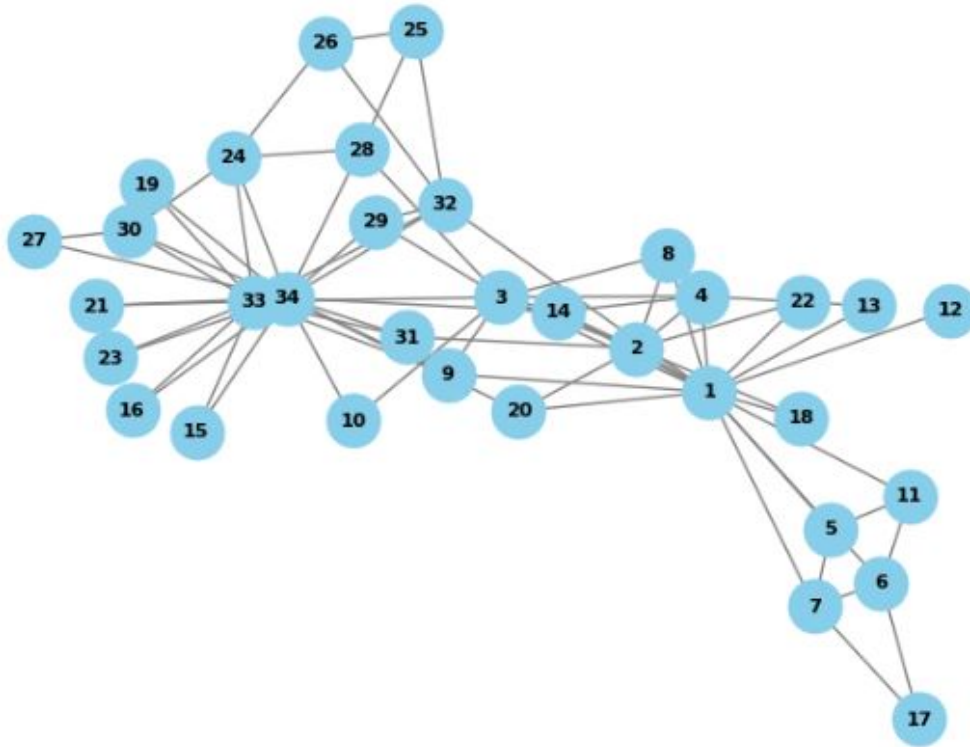


FIGURE 2 : Graphe du réseau des membres du club de karaté

Nous appliquons par la suite les codes développés dans la première partie pour calculer, pour ce réseau, la centralité du PageRank pour  $\alpha = 0.85$  et  $\beta = 1$ , la centralité d'intermédiarité et la centralité de proximité (voir fichier .ipynb).

Afin de résumer les résultats trouvés dans cette partie, nous présentons la table de synthèse de ce réseau :



Caractéristiques	Valeur	Noeud concerné
Nombre de noeuds	34	N/A
Nombre d'arêtes	78	N/A
Degré - min	1	12
Degré - max	17	34
Degré - moyen	4.67	N/A
Centralité (PageRank) - min	2.168008978217383	12
Centralité (PageRank) - max	22.875014662061847	34
Centralité (PageRank) - moyenne	6.666666666666667	N/A
Centralité d'intermédarité - min	0.0	8, 12, 13, 15, 16, 17, 18, 19, 21, 22, 23, 27
Centralité d'intermédarité - max	0.3553571428571464	1
Centralité d'intermédarité - moyenne	0.03753752290057451	N/A
Centralité de proximité - min	0.2931034482758621	17
Centralité de proximité - max	0.5862068965517242	1
Centralité de proximité - moyenne	0.4394032578030241	N/A

TABLE 8 : table de synthèse du réseau des membres du club de karaté

## 5 Troisième réseau : exemple d'un réseau de type Facebook

Nous commençons cette partie par visualiser, à l'aide de python, le graphe du réseau de type Facebook.

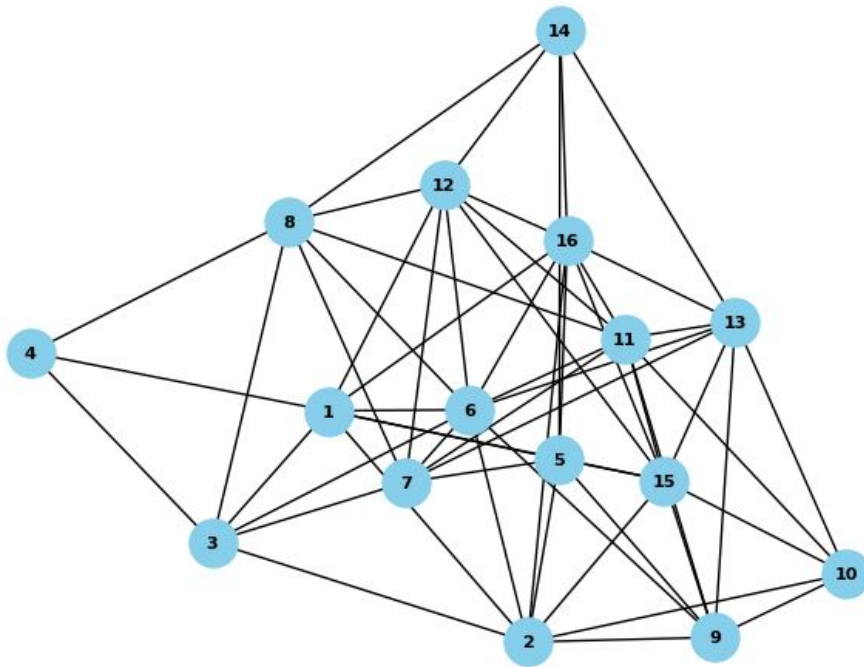


FIGURE 3 : graphe du réseau de type Facebook

Nous appliquons par la suite les codes développés dans la première partie pour calculer, pour ce réseau, la centralité du PageRank pour  $\alpha = 0.85$  et  $\beta = 1$ , la centralité d'intermédiarité et la centralité de proximité (voir fichier .ipynb).

Afin de résumer les résultats trouvés dans cette partie, nous présentons la table de synthèse de ce réseau :

Caractéristiques	Valeur	Noeud concerné
Nombre de noeuds	16	N/A
Nombre d'arêtes	58	N/A
Degré - min	3	4
Degré - max	10	6
Degré - moyen	7.25	N/A
Centralité (PageRank) - min	3.4252191945902126	4
Centralité (PageRank) - max	8.790562609389502	6
Centralité (PageRank) - moyenne	6.666666666666666	N/A
Centralité d'intermédiarité - min	0.0031746031746031746	4
Centralité d'intermédiarité - max	0.07532879818594106	1
Centralité d'intermédiarité - moyenne	0.038153344671201815	N/A
Centralité de proximité - min	0.5333333333333333	4
Centralité de proximité - max	0.8	15
Centralité de proximité - moyenne	0.6962622160448246	N/A

TABLE 9 : table de synthèse du réseau de type Facebook

## 6 Quatrième réseau : Réseau d'un groupe de chercheurs-physiciens

A l'instar des parties précédentes, nous commençons cette partie par visualiser, à l'aide de python, le graphe du réseau du groupe des chercheurs-physiciens.



Caractéristiques	Valeur	Noeud concerné
Nombre de noeuds	241	N/A
Nombre d'arêtes	1098	N/A
Degré - min	1	212
Degré - max	28	127
Degré - moyen	7.659751037344399	N/A
Centralité (PageRank) - min	1.8359389015566148	212
Centralité (PageRank) - max	22.03510279054408	127
Centralité (PageRank) - moyenne	6.666666666666665	N/A
Centralité d'intermédarité - min	N/A	N/A
Centralité d'intermédarité - max	N/A	N/A
Centralité d'intermédarité - moyenne	N/A	N/A
Centralité de proximité - min	0.29285714285714287	212
Centralité de proximité - max	0.7272727272727273	127
Centralité de proximité - moyenne	0.44157761188158373	N/A

TABLE 10 : table de synthèse du réseau des chercheurs-physiciens

## 7 Étude qualitative des résultats obtenus et comparaison des réseaux

Dans cette partie, nous étudierons et comparons les différents réseaux en termes de taille, densité, et les valeurs moyennes des différentes centralités.

### 1. Taille du réseau :

- Facebook : 16 nœuds
- club des physiciens : 241 nœuds
- club de karaté : 34 nœuds

### 2. Densité du réseau (nombre d'arêtes par nœud) :

- Facebook : 3.625 arêtes par nœud (58 arêtes / 16 nœuds)
- club des physiciens : 3.831 arêtes par nœud (923 arêtes / 241 nœuds)
- club de karaté : 2.294 arêtes par nœud (78 arêtes / 34 nœuds)

### 3. Distribution des degrés :

- Le réseau de type Facebook a un degré moyen de 7.25 avec un minimum de 3 et un maximum de 10.
- le réseau des membres du club des physiciens a un degré moyen de 7.66

avec un minimum de 1 et un maximum de 28.

- Le réseau des membres du groupe de karaté a un degré moyen de 4.59 avec un minimum de 1 et un maximum de 17.

#### 4. Centralité (PageRank) :

- Les valeurs de centralité du PageRank varient d'un réseau à l'autre, mais, remarquablement, tous les réseaux ont une moyenne de 6.67.

#### 5. Centralité d'intermédiarité :

Les deux réseaux, celui des membres du groupe de karaté et celui de type Facebook ont, au centième près, la même valeur moyenne de la centralité d'intermédiarité. Par contre, pour le dernier réseau, celui des physiciens, l'algorithme développé n'arrive pas à calculer les valeurs de la centralité d'intermédiarité puisque le réseau est divisé en quatre sous-réseaux disjoints, et donc, certains noeuds ne sont pas connectés et l'algorithme affiche une erreur de division par zéro.

#### 6. Centralité de proximité :

Le réseau de type Facebook représente la valeur moyenne de centralité de proximité la plus élevée, suivi par celui des chercheurs physiciens et enfin les membres du groupe de Karaté.