Utilisation des algorithmes TOPSIS, W-Topsis et K-Means pour l'identification des nœuds influents dans un réseau compliqué.

Master Sciences de données

Année universitaire 2022/2023

Réalisé par :

JALLOULI CHAIMAE

AZNAG ACHIK

CHARAF OUALID

Encadré par :

QAFFOU ISSAM

Table des matières

Introduction générale 4

Chapitre I. Contexte général 5

Introduction 5

Réseaux complexes 5

Graphes : outils de modélisation 5

Détection des nœuds influents 6

Chapitre II. Outils de détection des nœuds influents 7

1. Méthodes de détection des nœuds influents 7

Mesures de centralité 7

L'algorithme TOPSIS 8

L'algorithme K-Means 10

2. Méthodes comparatives 11

Modèle SI (Susceptible/Infected) 11

Chapitre III. Implémentation des algorithmes TOPSIS et w-TOPSIS pour la détection des nœuds influents 12

Introduction 12

1. Ensemble de données 12

2. Outils techniques 12

3. Implémentation de l'algorithme TOPSIS 12

Etapes 12

Utilisation du modèle comparatif SI 13

4. Implémentation de l'algorithme w-TOPSIS 14

Etapes 14

Comparaison entre topsis et w-topsis 14

5. Applications sur autres DataSets 15

Football Dataset 16

Zachary Karate Club Dataset 17

Conclusion 17

Chapitre IV. Utilisation de l'algorithme K-Means et Scores pour la détection des nœuds influents 19

1. K-Means : Partie I 19

Etapes 19

Comparaison 20

2. K-Means : Partie II 21

Etapes 21

Comparaison 21

3. K-Means sans initialisation 21

Etapes 21

Comparaison 21
4. Utilisation des Scores 22
Etapes 22
Remarque 22
Conclusion générale 24

Introduction générale

Avec le développement rapide des technologies de l'information, l'échelle des réseaux complexes augmente, ce qui rend la propagation des maladies et des rumeurs plus difficiles à contrôler. L'identification efficace et précise des nœuds influents est essentielle pour prédire et contrôler le système de réseau de manière pertinente.

L'identification des nœuds influents dans un réseau réel est un domaine de recherche actif, vaste et riche d'applications. Marketing viral, propagation de virus, confinement des rumeurs sont parmi les applications les plus connues.

Les nœuds et les bords de différents types de réseaux jouent divers rôles dans la structure et la fonction du réseau. Ces réseaux sont hétérogènes aux échelles macro, méso et micro.

L'un des problèmes les plus difficiles est l'identification des nœuds influents dans les réseaux sociaux dynamiques qui a attiré une attention croissante ces dernières années. Les vrais réseaux sociaux comme Facebook ou Twitter sont très changeants donc ils sont représentés comme des graphiques temporels qui évoluent dans le temps.

Les mesures de centralité sont des méthodes bien connues utilisées pour quantifier l'influence des nœuds en extrayant des informations de la structure du réseau. L'écueil de ces mesures est de repérer des nœuds situés à proximité les uns des autres, saturant leur zone d'influence commune. Néanmoins, ces mesures restent limitées et donnent des résultats moins performants par rapport aux autres algorithmes plus avancé.

Dans ce projet, nous allons entamer cette identification en utilisant des algorithmes notamment TOPSIS, WTOPSIS et AHP. Tout en introduisant les méthodes et outils implémentés, présentant les différentes étapes suivies et démontrant la mise en œuvre du projet.

Chapitre I. Contexte général

Introduction

Dans ce chapitre, nous allons effectuer une étude théorique, qui va servir comme base nécessaire pour comprendre le reste du projet. Cette étude théorique concerne la présentation des réseaux complexes, l'utilisation de la théorie des graphes et finalement la détection des nœuds influents.

Réseaux complexes

Dans le contexte de la théorie des réseaux, un réseau complexe est un graphe (réseau) avec des caractéristiques qui ne se produisent pas dans des réseaux simples tels que des treillis ou des graphes aléatoires mais se produisent souvent dans des réseaux représentant des systèmes réels.

L'étude des réseaux complexes est un domaine de recherche scientifique jeune et actif (depuis 2000) inspiré en grande partie par les découvertes empiriques de réseaux du monde réel tels que les réseaux informatiques, les réseaux biologiques, les réseaux technologiques, les réseaux cérébraux, réseaux climatiques et réseaux sociaux.

Graphes : outils de modélisation

Un graphe est une collection d'éléments mis en relation entre eux . Géométriquement, on représente ces éléments par des points (les sommets) reliés entre eux par des arcs de courbe (les arêtes). Selon que l'on choisit d'orienter les arêtes ou de leur attribuer un poids (un coût de passage), on parle de graphes orientés ou de graphes pondérés.

La théorie des graphes s'intéresse à leurs multiples propriétés : existence de chemins les plus cours, de cycles particuliers, nombre d'intersections dans le plan, problèmes de coloriage…

Aujourd'hui, les graphes trouvent plusieurs applications dans la modélisation des réseaux (routiers, informatiques, etc..). Par ailleurs, la théorie des graphes a fourni des problèmes algorithmiques cruciaux en théorie de la complexité.

Détection des nœuds influents

Dans les réseaux, tous les nœuds n'ont pas la même importance, et certains sont plus importants que d'autres. La question de trouver les nœuds les plus importants dans les réseaux a été largement abordée, en particulier pour les nœuds dont l'importance est liée à la connectivité du réseau.

Ces nœuds sont généralement appelés nœuds critiques. Le problème de détection de nœud critique (CNDP) est le problème d'optimisation qui consiste à trouver l'ensemble de nœuds dont la suppression dégrade au maximum la connectivité du réseau selon certaines métriques de connectivité prédéfinies. Prédéfinie, différentes variantes ont été développées.

Chapitre II. Outils de détection des nœuds influents

Dans ce chapitre, nous allons présenter les différentes méthodes de détection des nœuds influents accompagnés des algorithmes d'évaluation, passant du plus simple au plus avancé.

1. Méthodes de détection des nœuds influents

Mesures de centralité

Les mesures de centralité sont un outil essentiel pour comprendre les réseaux, souvent aussi appelés graphes.

Ces algorithmes utilisent la théorie des graphes pour calculer l'importance d'un nœud donné dans un réseau. Ils coupent les données bruyantes, révélant les parties du réseau qui nécessitent une attention, mais ils fonctionnent tous différemment. Chaque mesure a sa propre définition de "l'importance", vous devez donc comprendre leur fonctionnement pour trouver la meilleure pour vos applications de visualisation de graphiques.

- Degree Centrality : Attribue un score d'importance basé simplement sur le nombre de liens détenus par chaque nœud. Cette mesure nous montre le nombre de connexions directes chaque nœud a-t-il avec d'autres nœuds du réseau. Elle est souvent utilisée pour trouver des personnes très connectées, des personnes populaires, des personnes susceptibles de détenir la plupart des informations ou des personnes pouvant se connecter rapidement au réseau plus large.

- Betweenness Centrality : Mesure le nombre de fois qu'un nœud se trouve sur le chemin le plus court entre d'autres nœuds. Cette mesure montre quels nœuds sont des « ponts » entre les nœuds d'un réseau. Pour ce faire, il identifie tous les chemins les plus courts, puis compte le nombre de fois où chaque nœud tombe sur un. Cette mesure est utilisée Pour trouver les individus qui influencent le flux autour d'un système.

- Closeness Centrality : Note chaque nœud en fonction de sa « proximité » avec tous les autres nœuds du réseau. Cette mesure calcule les chemins les plus courts entre tous les nœuds, puis attribue à chaque nœud un score basé sur sa somme des chemins les plus courts, et elle est utilisée pour trouver les individus les mieux placés pour influencer le plus rapidement l'ensemble du réseau.

- Eigenvector Centrality : Comme Degree Centrality, EigenCentrality mesure l'influence d'un nœud en fonction du nombre de liens qu'il a avec d'autres nœuds du réseau. EigenCentrality va ensuite un peu plus loin en prenant également en compte le niveau de connexion d'un nœud, le nombre de liens de ses connexions, etc. à travers le réseau. Cette mesure calcule les connexions étendues d'un nœud,

EigenCentrality peut identifier les nœuds ayant une influence sur l'ensemble du réseau, pas seulement ceux qui y sont directement connectés. Elle est pratique pour comprendre les réseaux sociaux humains, mais aussi pour comprendre les réseaux comme la propagation des logiciels malveillants.

L'algorithme TOPSIS

TOPSIS, connue sous le nom de Technique for Order of Preference by Similarity to Ideal Solution, est une méthode d'analyse décisionnelle multicritères. Il compare un ensemble d'alternatives sur la base d'un critère prédéfini. La méthode est utilisée dans l'entreprise dans diverses industries, chaque fois que nous devons prendre une décision analytique basée sur les données collectées.

La logique de TOPSIS est basée sur le concept que l'alternative choisie doit avoir la distance géométrique la plus courte de la meilleure solution et la distance géométrique la plus longue de la pire solution.

Une telle méthodologie permet de trouver des compromis entre les critères lorsqu'une mauvaise performance sur l'un peut être annulée par une bonne performance sur un autre critère. Cela fournit une forme de modélisation assez complète car nous n'excluons pas des solutions alternatives basées sur des seuils prédéfinis.

- Etapes de l'algorithme Topais

i- Création d'une matrice d'évaluation de M alternatives et N critères

$$(a_{ij})_{MxN}$$

Dans notre exemple, les alternatives seront les différents nœuds du réseau, et les critères seront les mesures de centralité.

ii- Normalisation de la matrice d'évaluation

$$\alpha_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{M}(a_{ij})^2}}$$

iii- Calcul du WeightedNormalized Matrix

$$\chi_{ij} = \alpha_{ij} * \omega_j$$

$$\sum_{j=1}^{N} \omega_j = 1$$

iv- Déterminer le meilleur et le pire alternative pour chaque critère

$$\chi_j^b = max\chi_{ij}$$

$$\chi_j^w = min\chi_{ij}$$

v- Calculer la distance euclidienne entre l'alternative et meilleur/pire alternative

$$d_j^b = \sqrt{\sum_{j=1}^{N}(\chi_j^i - \chi_j^b)^2}$$

$$d_j^w = \sqrt{\sum_{j=1}^{N} (\chi_j^i - \chi_j^w)^2}$$

vi- Calculer la similarité entre chaque alternative et le pire alternative

$$S_i = \frac{d_i^w}{d_i^w + d_i^b}$$

vii- Trier les alternatives selon la valeur du Topsis dans l'ordre décroissant.

De cette manière, on obtient un ensemble d'alternatives ordonnés selon des critères spécifiques.

L'algorithme K-Means

- Définition

Un algorithme non supervisé de clustering non hiérarchique. Il permet de regrouper en clusters distincts les observations du data set. Ainsi les données similaires se retrouveront dans un même cluster. Par ailleurs, une observation ne peut se retrouver que dans un cluster à la fois (exclusivité d'appartenance). Une même observation, ne pourra donc, appartenir à deux clusters différents.

- Notion de similarité

Pour pouvoir regrouper un jeu de données en cluster distincts, l'algorithme K-Means a besoin d'un moyen de comparer le degré de similarité entre les différentes observations. Ainsi, deux données qui se ressemblent, auront une distance de dissimilarité réduite, alors que deux objets différents auront une distance de séparation plus grande.

- Calcul de distance

Les littératures mathématiques et statistiques regorgent de définitions de distance, les plus connues pour les cas de clustering sont :

La distance Euclidienne

C'est la distance géométrique qu'on apprend au collège. Soit une matrice à variables quantitatives. Dans l'espace vectoriel . La distance euclidienne entre deux observations et se calcule comme suit :

$$X = \sqrt{(X_b - X_a)^2 + (Y_b - Y_a)^2}$$

La distance de Manhattan

La distance entre deux points parcourus par un taxi lorsqu'il se déplace dans une ville où les rues sont agencées selon un réseau ou un quadrillage. Un taxi-chemin est le trajet fait par un taxi lorsqu'il se déplace d'un nœud du réseau à un autre en utilisant les déplacements horizontaux et verticaux du réseau.

- Fonctionnement de l'algorithme K-Means

K-means est un algorithme itératif qui minimise la somme des distances entre chaque individu et le centroïde. Le choix initial des centroïdes conditionne le résultat final.

Admettant un nuage d'un ensemble de points, K-Means change les points de chaque cluster jusqu'à ce que la somme ne puisse plus diminuer. Le résultat est un ensemble de clusters compacts et clairement séparés, sous réserve de choisir la bonne valeur du nombre de clusters.

2. Méthodes comparatives

Modèle SI (Susceptible/Infected)

Le modèle le plus simple d'un nœud infectieux catégorise les nœuds comme susceptibles ou infectieux (SI). On peut imaginer que les nœuds sensibles sont en bonne santé et que les nœuds infectieux sont malades.

Un nœud sensible peut devenir contagieux par contact avec un infectieux. Ici, et dans tous les modèles ultérieurs, nous supposons que la population étudiée est bien mélangée, de sorte que chaque nœud a une probabilité égale d'entrer en contact avec tous les autres nœuds.

Chapitre III. Implémentation des algorithmes TOPSIS et w-TOPSIS pour la détection des nœuds influents

Introduction

Dans cette partie, on attaque la partie pratique du projet. Comme première étape, nous avons implémenté l'algorithme TOPSIS accompagné des 4 mesures de centralité : Degree Centrality, Betweenness Centrality, Closeness Centrality et Eigenvector Centrality.

1. Ensemble de données

Pour ce projet, nous avons choisi d'utiliser le dataset « Ego-Facebook ». Ce réseau dirigé contient 2900 nœuds qui représentent des amitiés utilisateur-utilisateur de Facebook. Un nœud représente un utilisateur. Un bord indique que l'utilisateur représenté par le nœud de gauche est un ami de l'utilisateur représenté par le nœud de droite.

2. Outils techniques

- Langage de programmation : Python
- Libraires implémentées

  - Numpy
  - Pandas
  - MatplotLib
  - NetworkX
  - Math
  - …

3. Implémentation de l'algorithme TOPSIS

Etapes

Dans cette partie, nous avons appliqué l'algorithme TOPSIS sur notre ensembles de données, en suivant les étapes mentionnées dans la partie théorique, pour obtenir une matrice d'évaluation.

On applique l'algorithme Topsis sur la matrice d'évaluation, on la donnant comme argument, ainsi que la liste des poids [0.2, 0.3,0.2,0.3] qui a donné un résultat optimal.

Par la suite, on passe à calculer la valeur du Closeness, sur laquelle on va se baser pour trier nos nœuds du plus au moins influent.

Notre résultat sera diffusé sous forme d'un tableau, contenant pour chaque nœud, les quatre mesures de centralité ainsi que la valeur du Closeness.

```
[ ] result = pd.concat([closeness_topsis, centralities_closeness], axis=1)
    result.head(10)
```

	S+	S-	C	node2	node	dc	bc	cc	ec	C
0	0.140349	0.061395	0.304321	107	107	0.258791	0.480518	0.459699	2.606940e-04	0.913277
1	0.201098	0.001590	0.007846	1684	1684	0.196137	0.337797	0.393606	7.164260e-06	0.695566
2	0.201175	0.001467	0.007238	1912	1912	0.186974	0.229295	0.350947	9.540696e-02	0.496063
3	0.201098	0.001590	0.007846	3437	3437	0.135463	0.236115	0.314413	9.531613e-08	0.488865
4	0.201175	0.001467	0.007238	0	0	0.085934	0.146306	0.353343	3.391796e-05	0.304321
5	0.201142	0.001513	0.007464	1085	1085	0.016345	0.149015	0.357852	3.164082e-06	0.297379
6	0.201219	0.001424	0.007029	698	698	0.016840	0.115330	0.271189	1.116876e-09	0.231106
7	0.200996	0.001857	0.009154	567	567	0.015602	0.096310	0.328881	9.932295e-06	0.193646
8	0.201197	0.001443	0.007120	58	58	0.002972	0.084360	0.397402	5.898120e-04	0.169464
9	0.200665	0.002947	0.014476	428	428	0.028479	0.064309	0.394837	5.990065e-04	0.132923

Figure 1: cocnaténation des dataframe

Utilisation du modèle comparatif SI

A l'aide de l'algorithme SI, nous avons arrivé à visualiser les résultats obtenus avec Topsis sous forme de courbes, et les comparer avec les mesures de centralité.

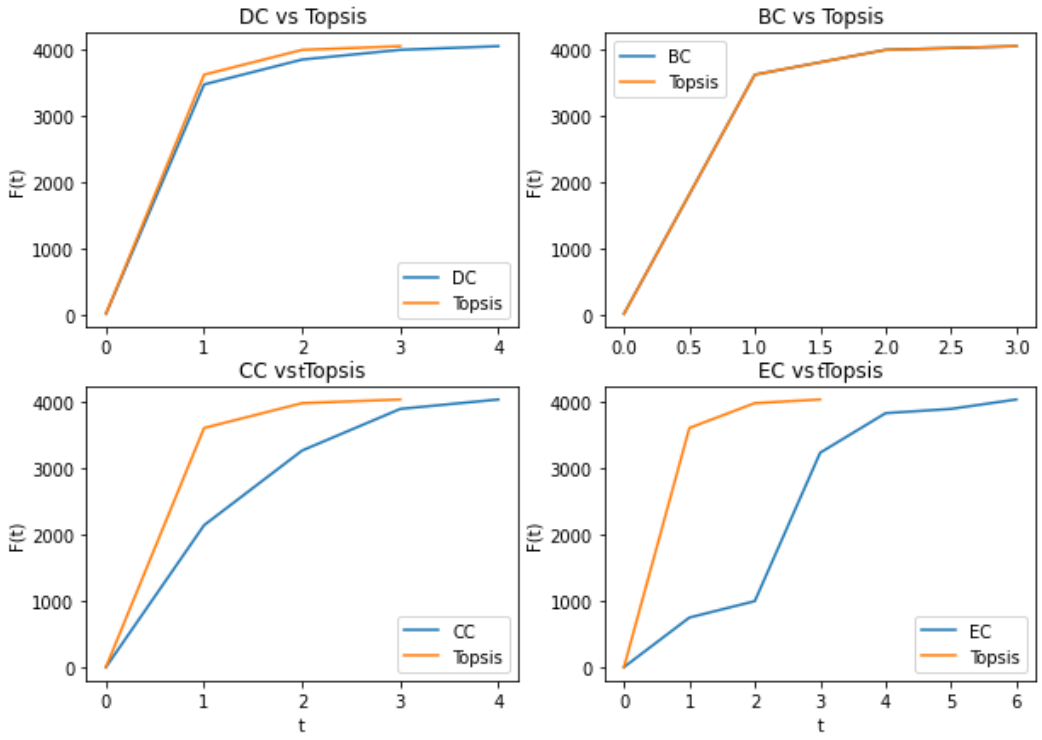


Figure 2: comparaison entre topsis et mesures de centralité

4. Implémentation de l'algorithme w-TOPSIS

Etapes

Pour appliquer cet algorithme, nous avons principalement passé par les mêmes étapes de l'algorithme TOPSIS régulier, c'est-à-dire :

- L'acquisition et génération des données
- Normalisation de la donnée
- Détermination du poids
- Sélection des meilleures alternatives en utilisant TOPSIS

Au niveau de l'algorithme w-TOPSIS, une étape supplémentaire s'ajoute afin de calculer des poids à valeurs plus optimales.

On calcule la valeur de l'entropie relativement aux éléments de la matrice d'évaluation en effectuant des calculs purement mathématiques, et on utilise ces valeurs d'entropie pour calculer les poids.

Finalement, on passe par les mêmes étapes utilisées au niveau de Topsis pour trier les nœuds du plus au moins infulent.

Comparaison entre topsis et w-topsis

Vu la grande similarité entre les deux algorithmes Topsis et W-Topsis, il est évident d'effectuer entre les deux, prenant en considération que le premier

fonctionne d'une manière statique (valeurs des poids manuels), alors que le deuxième est plus dynamique (calcul des poids à l'aide de l'entropie).

Dans la figure ci-dessous, on représente au niveau de quatre courbes Topsis, W-Topsis et à chaque fois une mesure de centralité afin de pouvoir les comparer.

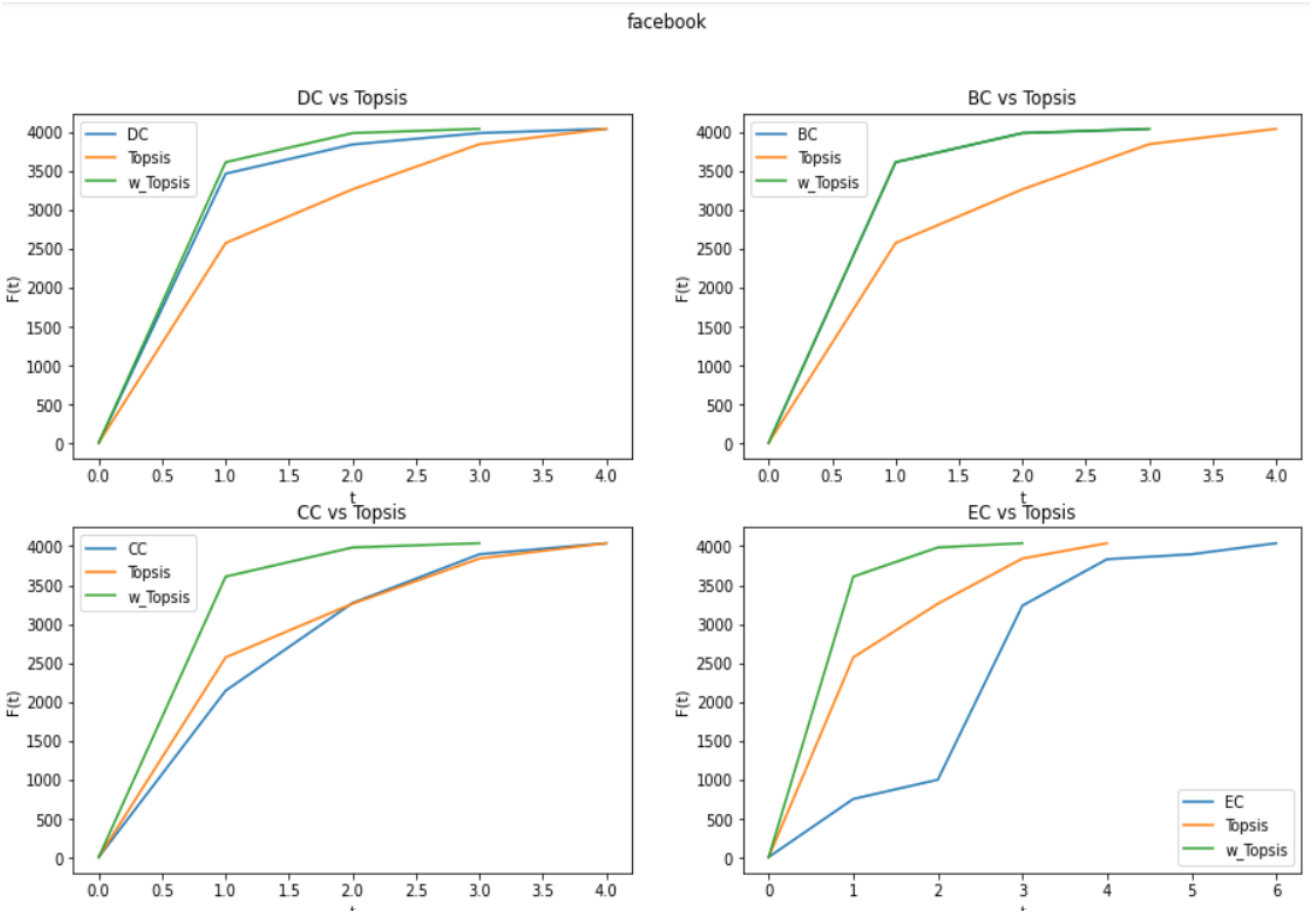


Figure 3: comparaison entre Topsis, W-Topsis et les mesures de centalité

Remarque

On constate que la courbe de w-topsis est au-dessus des autres presque dans tous les graphes.

Conclusion

On déduit que les poids calculés par la méthode d'entropie étaient optimaux et par conséquent la méthodes w-topsis est plus efficace grâce à sa dynamicité, par rapport à la méthode topsis.

5. Applications sur autres DataSets

Après avoir appliqué les 2 algorithmes sur le dataset « Facebook », et après avoir aboutit aux conclusions après la comparaison entre eux, il est

conseillé d'appliquer la même démarche en utilisant cette fois-ci d'autre jeux de données.

Football Dataset

Cet ensemble de données a été publié en 1998 par L. Krempel. Il est décrit comme un graphe orienté pondéré, ou un réseau pondéré, contenant dans sa totalité 115 nœuds.

La figure ci-dessous montre la comparaison entre Topsis, W-Topsis et les quatre mesures de centralité.

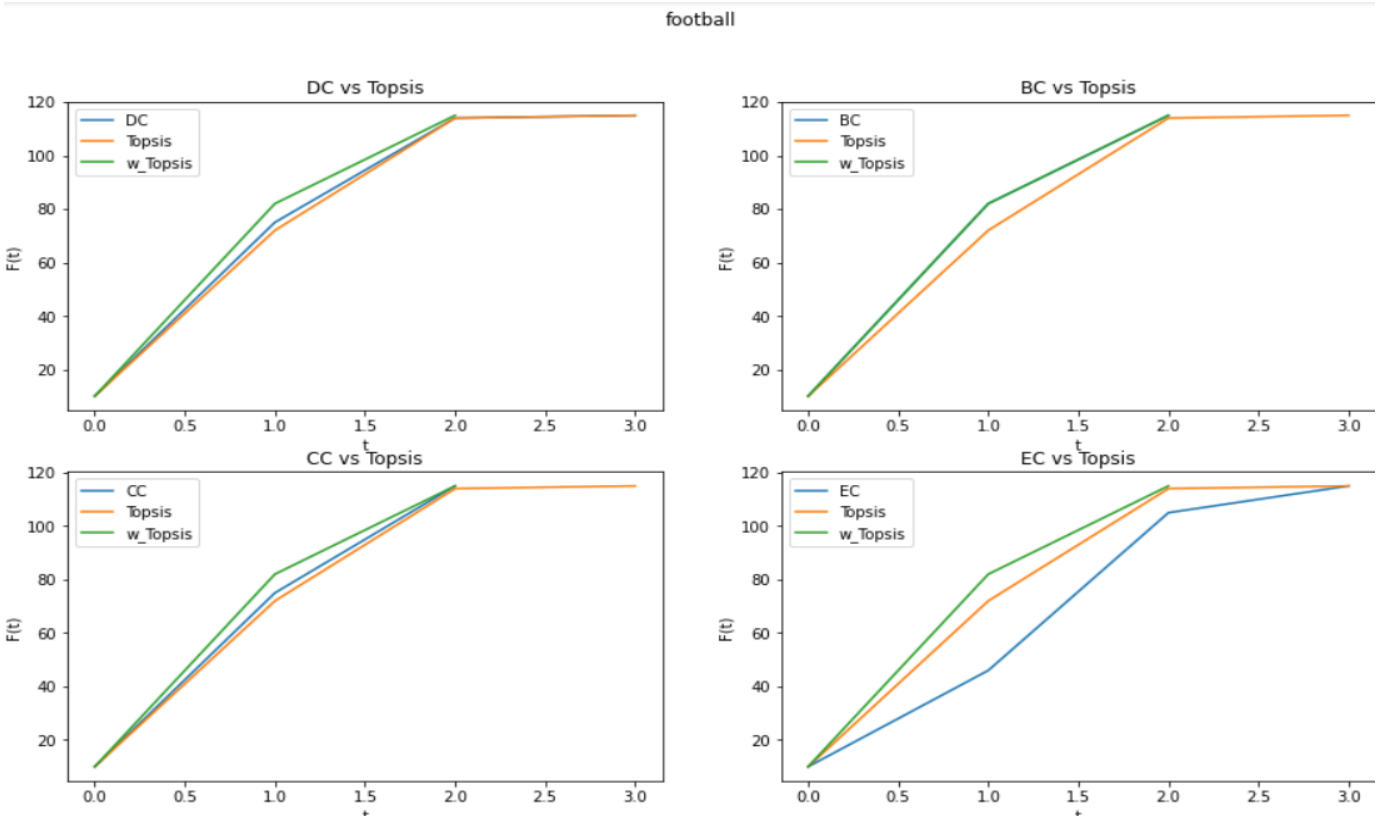


Figure 4: Comparaison pour le Dataset Football

Zachary Karate Club Dataset

Les données ont été recueillies auprès des membres d'un club de karaté universitaire par Wayne Zachary en 1977. Chaque nœud représente un membre du club et chaque arête représente une égalité entre deux membres du club. Le réseau n'est pas orienté.

La figure ci-dessous montre la comparaison entre Topsis, W-Topsis et les quatre mesures de centralité.

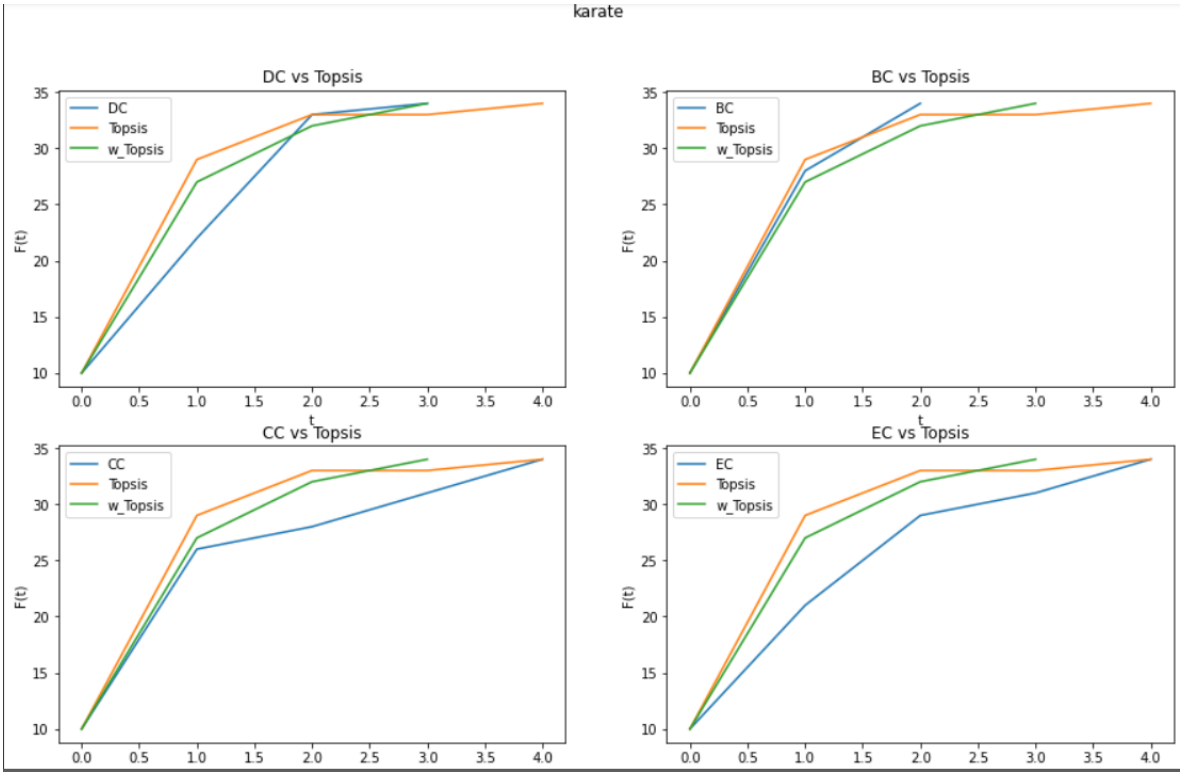


Figure 5: Comparaison pour le DatasetKarate

Conclusion

Dans la figure ci-dessous, on présente une visualisation des graphes présents au niveau de chacun des ensembles des données qu'on a utilisé.

Pour le premier ensemble, on remarque qu'on a un graphe géant.

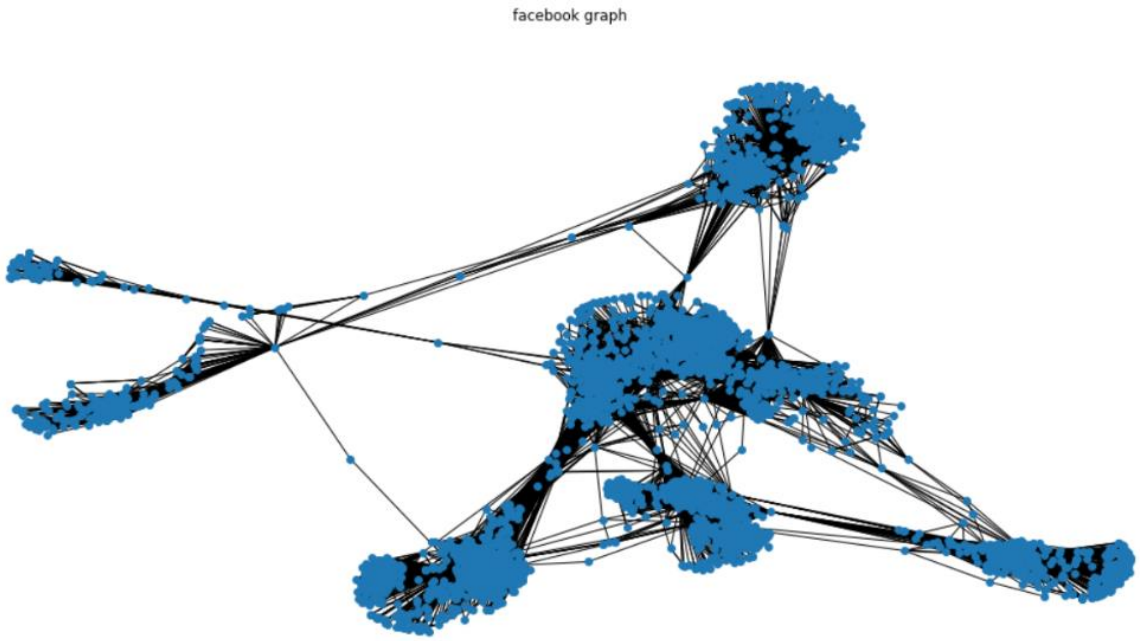


Figure 6: Visualisation du graphe Facebook

Pour le deuxième ensemble, on a un graphe à un nombre de nœuds moyen.

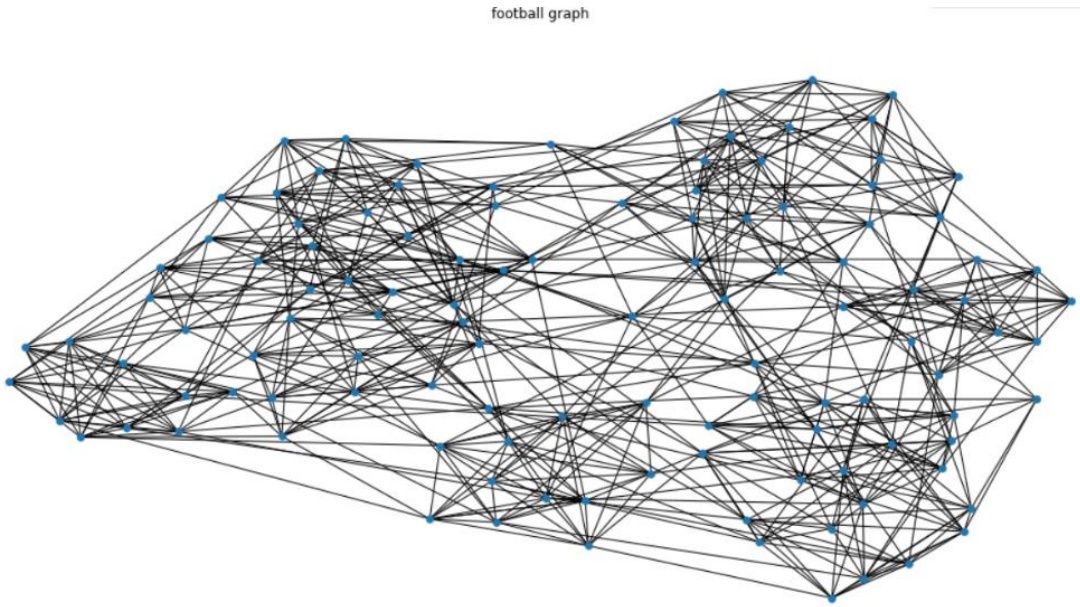


Figure 7: Visualisation du graphe Football

Et finalement, pour le troisième ensemble, il s'agit d'un petit graph à un nombre limité de nœuds.

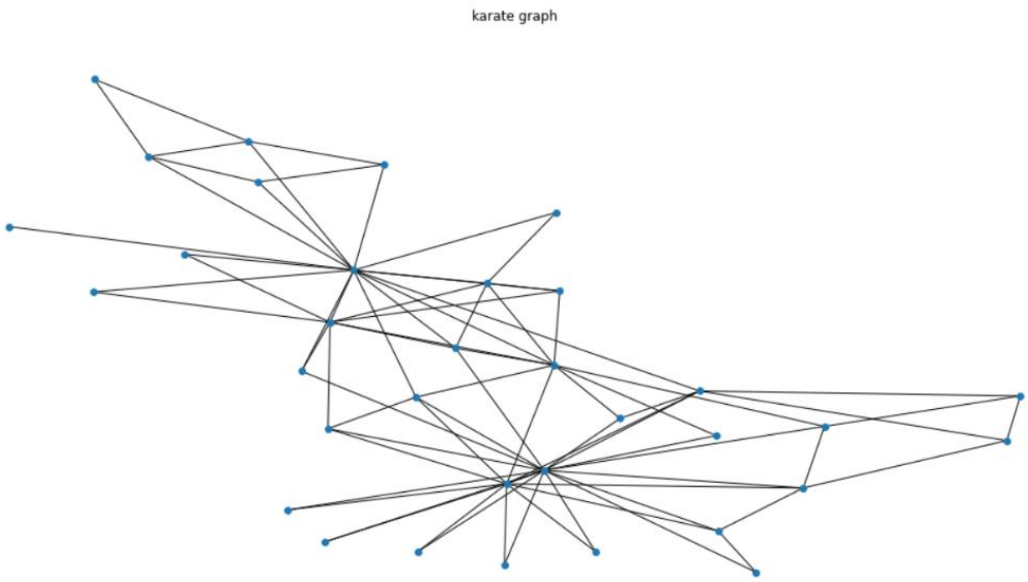


Figure 8: Visualisation du graphe Karate

La visualisation de la taille de chaque graphe, et lier cette dernière aux résultats obtenus après avoir appliqué les algorithmes nous mènent vers la conclusion que le comportement de Topsis et W-Topsis, dépend de la taille du réseau. Du coup, W-Topsis donne une performance meilleure lorsqu'il est associé à un graphe plus grand.

Chapitre IV. Utilisation de l'algorithme K-Means et Scores pour la détection des nœuds influents

1. K-Means : Partie I

Etapes

Dans cette étape, on utilise l'algorithme K-Means pour créer des clusters du graph entier, et comparer leurs centres avec les autres critères utilisés au niveau des étapes précédentes, et conclure qui a donné un Ranking meilleur.

Dans un premier lieu, on commence par calculer les k nœuds les plus influents de la manière qu'on a fait dans la partie précédente, avec k un nombre choisi par l'utilisateur, et ça en calculant les mesures de centralité et en créant la matrice d'évaluation.

Dans cet exemple, l'utilisateur a saisi le nombre 10 comme valeur de la variable k, ce qui nous donne 10 nœuds influents du réseau.

Au niveau de l'algorithme K-Means, on va spécifier 2 paramètres principaux :

- Numéro de clusters désiré, dans notre cas il s'agit du même nombre saisi par l'utilisateur.
- Les centres que l'algorithme va utiliser comme initialisation, dans notre cas c'est le fichier CSV qu'on a importé.

On va passer ses paramètres, en utilisant les mesures de centralité comme cordonnées et entraîner le modèle.

Après avoir obtenir les clusters et leurs centres à l'aide de K-Means, on effectue un calcul de distance entre les mesures de centralité de ces centres et les nœuds du graphe, pour obtenir d'eux le nœud le plus proche.

Comparaison

Après avoir calculé les k nœuds les plus influents, et après avoir crée k clusters et optimiser leurs centres à l'aide de la distance, c'est temps d'appliquer l'algorithme SI pour effectuer une étude comparative entre les critères suivants :

- Les mesures de centralité
- Le résultat obtenu par l'algorithme W-Topsis
- Les centres des clusters effectués par K-Means

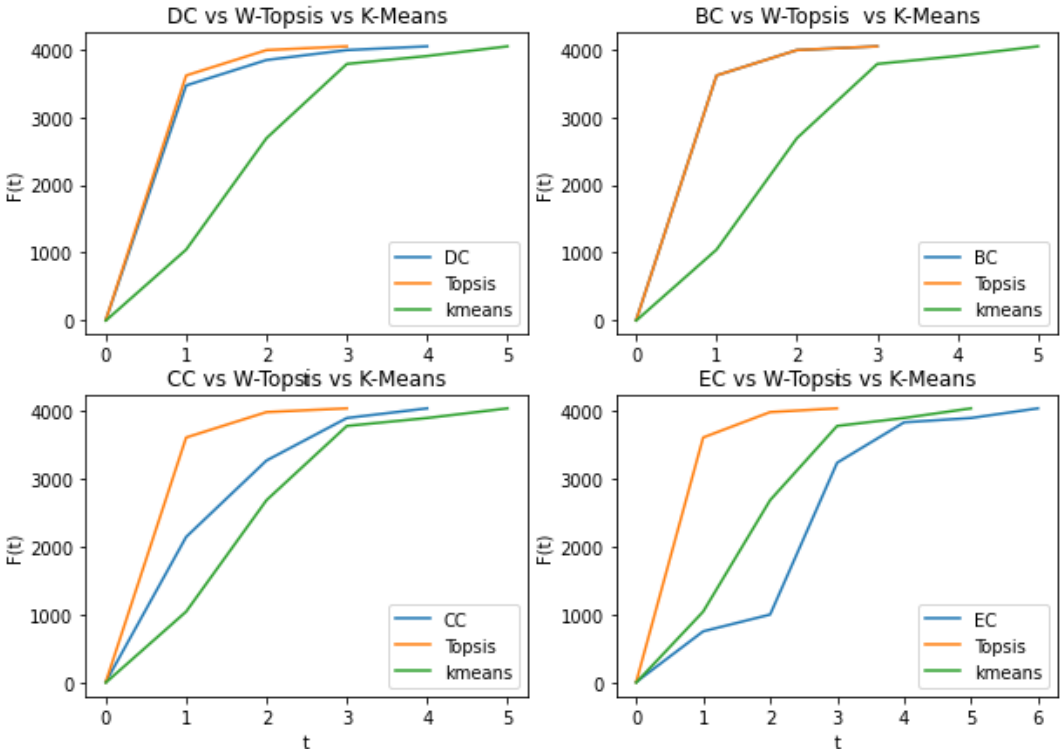


Figure 9: Comparaison entre K-Means et W-Topsis (partie I)

D'une manière générale, l'utilisation de l'algorithme K-Means, donne des résultats assez faibles en comparaison avec les mesures de centralité et l'algorithme W-Topsis, à l'exception de la masure Eigenvector Centrality.

2. K-Means : Partie II

Etapes

Dans cette partie, les clusters deviennent eux-mêmes des sous-graphes, sur lesquels on va appliquer l'algorithme W-Topsis, obtenir les k plus influents de chaque sous-graphe, prendre le meilleur, combiner les 10 meilleurs des graphes, et les comparer.

Comparaison

Sur un cluster choisi, on applique l'algorithme SI et on arrive à visualiser les courbes

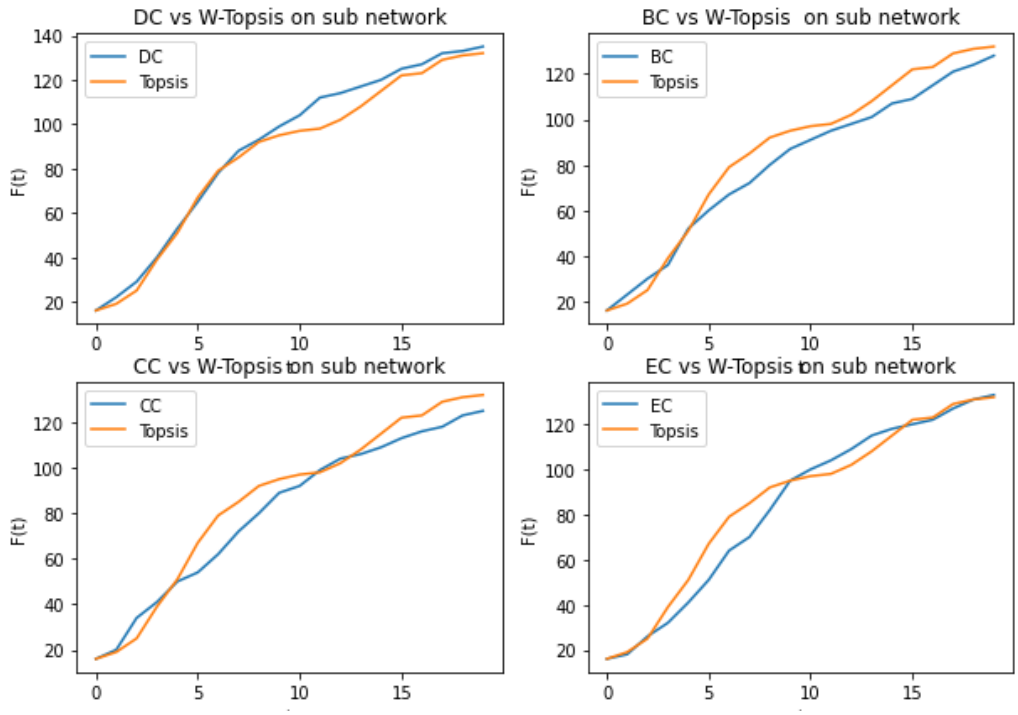


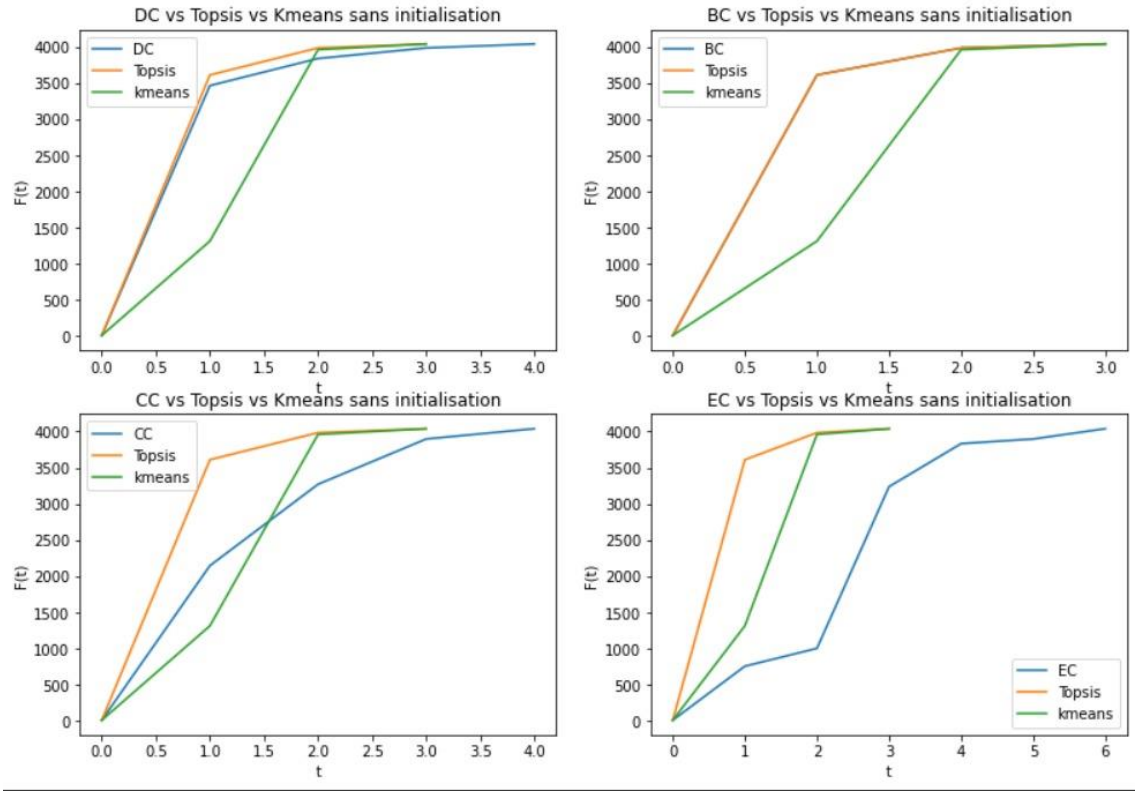
Figure 10 : Comparaison entre K-Means et W-Topsis (partie II)

3. K-Means sans initialisation

Etapes

Dans cette partie, nous avons implémenté l'algorithme K-Means sans initialisation.

Comparaison



En utilisant l'algorithme SI, on remarque que Topsis reste toujours le meilleur par rapport à K-Means et aux mesures de centralité.

4. Utilisation des Scores

Etapes

A cette étape là, chaque nœud est caractérisé par 5 cordonnées principaux : les quatres mesures de centralité ainsi que la valeur du Closeness.

Une autre méthode consiste à attribuer à chaque nœud un nouvel attribut, qu'on appellera le Score.

Le calcul du Score pour chaque nœud du réseau va se faire selon la formule suivante :

$$0.5 * BC + 0.3 * DC + 0.2 * CC$$

On va appliquer l'algorithme K-Means, cette fois-ci en se basant sur le Score de chaque nœud, et on termine par appliquer W-Topsis sur les clusters obtenus.

Remarque

L’utilisation des scores a donné des clusters erronés, même si combiné avec l’algorithme Topsis. A titre d’exemple, un seul cluster contient 4027 nœuds, ce qui n’est pas un résultat désirable. Donc, on n’a pas pris cette méthode en considération.

Conclusion générale

La détection des nœuds influents est une discipline très intéressante, qui peut être utile dans les différents domaines de notre quotidien, c'est pour cette raison que plusieurs algorithme ont été fait pour donner le meilleur résultat possible.

Dans ce projet, nous avons essayé quelques méthodes pour pouvoir extraire les nœuds les plus influents dans un réseau donné. Nous avons commencé par les mesures de centralité qui ont donné un résultat satisfaisant, mais qui a été facilement surpassé par l'algorithme TOPSIS, qui est basé sur des calculs mathématiques plus précis.

Cet algorithme aussi n'est pas optimal, vu qu'il est statique, et nécessite un choix manuel des poids utilisés pour calculer la matrice pondérée. Ce choix, qui est souvent aléatoire, et nécessite plusieurs tentatives, nous a mené a essayé un algorithme similaire.

W-Topsis, ou Weighted Topsis, comme son nom l'indique, est une version améliorée, plus dynamique de l'algorithme Topsis, qui utilise le calcul de l'entropie pour générer les poids de la matrice, et qui donne des résultats meilleurs.

Les résultats de W-Topsis restent les meilleurs même avec l'utilisation de la méthode de Clustering avec K-Means et l'introduction des scores.

Vers la fin, ce projet a été une introduction vers les graphes et les nœuds influents, en passant par une documentation théorique, passant par l'application des différents algorithmes et terminant par comparer entre les différentes méthodes.

Fin.