

# Comment analyser un réseau social en utilisant les graphes?

Un Apprentissage Par Problème (APP) destiné aux étudiants  
du module Algorithmique avancée 3



## Éléments de contexte :

Un réseau social est un ensemble d'utilisateurs interconnectés par des liens, qui représentent les relations et interactions entre les membres de ce réseau. Ces liens peuvent inclure des relations de communication, des interactions sociales telles que l'amitié, le suivi ou l'abonnement, des collaborations professionnelles, ou d'autres formes d'engagement et de connexion entre les utilisateurs.



## Travail demandé :

En tant qu'équipe d'ingénieurs, vous souhaitez réaliser une application d'analyse de réseaux sociaux. Cette application doit permettre d'importer un fichier contenant les données sur les interactions des utilisateurs, puis effectuer les tâches suivantes :

- Générer la matrice d'adjacence du réseau social (une matrice binaire) et la stocker dans un fichier texte.
- Utiliser la matrice d'adjacence générée pour identifier les meilleurs leaders d'opinion représentés par :



- Les nœuds ayant le maximum d'arêtes entrantes dans un graphe orienté « également appelé graphe dirigé ou direct ».
  - Les nœuds ayant le maximum de voisins (maximum d'arêtes) dans un graphe non-orienté « appelé aussi graphe non dirigé ou indirect ».
- 
- Pour chaque leader, afficher la liste de ses followers, représentés par les nœuds ayant des arêtes entrantes dans son nœud si le graphe est orienté ou par les nœuds voisins si le graphe est non-orienté.
  - Identifier et afficher les meilleurs followers du graphe social (les nœuds ayant le maximum d'arêtes sortantes vers les autres nœuds si le graphe est orienté).
  - Utiliser l'algorithme de parcours en largeur (BFS) pour déterminer le chemin de transmission de l'information.
  - Afficher le graphe du réseau social en utilisant la bibliothèque NetworkX.
  - Afficher les leaders avec une couleur distincte pour les mettre en évidence. Si les nœuds des leaders ne sont pas clairement visibles, effectuer plusieurs exécutions jusqu'à ce qu'ils le soient.

Note : Vous êtes autorisé à utiliser le package NetworkX uniquement pour l'affichage du graphe et non pour les autres aspects du problème.

## Les ressources pour traiter la situation-problème

### ➤ Les jeux de données :

#### 1. Fichier « Exemple » :

Utilisez le fichier texte **exemple.txt** pour :

- Générer la matrice d'adjacence qui représente le graphe orienté du fichier « exemple ».
- En utilisant cette matrice générée :
  - Identifier les deux leaders les plus importants dans le graphe en utilisant sa matrice d'adjacence.
  - Identifier les followers de chaque leader.
  - Identifier les 2 best followers du graphe.
- Identifier le chemin le plus court entre les deux leaders en utilisant la recherche en largeur.
- Dessiner le graphe.
- Afficher les leaders avec une couleur distincte.

#### 2. Fichier « Club » :

Utilisez le fichier texte **club.txt**, qui représente les relations entre les joueurs d'un club de karaté nommé Zachary, pour :



- Générer la matrice d'adjacence qui représente le graphe non-orienté du fichier « club ».
- En utilisant cette matrice générée :
  - Identifier les deux leaders les plus importants dans le graphe en utilisant sa matrice d'adjacence.
  - Identifier les followers de chaque leader.
- Identifier le chemin le plus court entre les deux leaders en utilisant la recherche en largeur.
- Dessiner le graphe.
- Afficher les leaders avec une couleur distincte.

### 3. Fichier « students » :

Le réseau social « *Students cooperation* » a été construit à partir des données recueillies lors d'un cours de "sécurité informatique et réseau", un cours obligatoire à l'université Ben-Gurion. Le réseau social contient des données recueillies auprès de 185 étudiants participants issus de deux départements différents. Le réseau social du cours a été créé en analysant la coopération implicite et explicite entre les étudiants pendant qu'ils faisaient leurs devoirs. Le graphe de coopération des étudiants contient 185 nœuds, 360 liens et 3 types de liens.

- Nombre de nœuds : 185
- Nombre d'arêtes : 360

Utilisez le lien suivant pour télécharger ce fichier:

<https://data4goodlab.github.io/dataset.html>

Utilisez le fichier **students.csv** pour :

- Générer la matrice d'adjacence qui représente le graphe orienté du fichier « students ».
- En utilisant cette matrice générée :
  - Identifier le meilleur leader dans le graphe en utilisant sa matrice d'adjacence.
  - Identifier les followers de ce leader.
  - Identifier les 5 best followers du graphe.
- Identifier le chemin le plus court entre les best followers en utilisant la recherche en largeur.

### 4. Fichier « Anybeat » :

Anybeat est une communauté en ligne, un lieu de rassemblement public où vous pouvez interagir avec des personnes de votre quartier ou du monde entier.

- Nombre de nœuds : 67,054
- Nombre d'arêtes : 12,645

Utilisez le lien suivant pour télécharger ce fichier:

<https://data4goodlab.github.io/dataset.html>



Utilisez le fichier **anybeatAnonymized.csv** pour :

- Générer la matrice d'adjacence qui représente le graphe orienté du fichier « students ».
- En utilisant cette matrice générée, analysez ce réseau social en identifiant les leaders, les followers et le chemin le plus important à tracer sur ce graphe.

## Objectifs d'apprentissage de l'APP:

- Maîtriser les notions de base des graphes (matrice d'adjacence, types de graphes, types d'arêtes, recherche en largeur...)
- Afficher les graphes