

UNIVERSITÉ CADI AYYAD
FACULTÉ DES SCIENCES SEMLALIA-MARRAKECH

Filière : Science des Données

Département : Informatique

Analyse des Réseaux Complexe et Détection des Communautés

Encadré par :

Pr. QUAFFOU

Réalisé par :

ENNAIR Aya

LAGHDAF Khaoula

ANNÉE UNIVERSITAIRE 2023-2024

Liste des figures

Figure 1 : Caractéristiques des réseaux	3
Figure 2 : K-Means Principe	8
Figure 3 : Girvan-Newman Principe	9
Figure 4 : Réseau du Facebook Dataset.....	12
Figure 5 : Réseau avec Visualisation des Clusters (K-Means).....	13
Figure 6 : Réseau avec Visualisation des Clusters (Louvain)	14
Figure 7 : Sous-Graphe Représentatif du Dataset Facebook.....	15
Figure 8 : Visualisation du Sous-Graphe Représentatif avec Gephi	16
Figure 9 : Sous-Réseau avec Visualisation des Clusters (K-Means).....	17
Figure 10 : Sous-Réseau avec Visualisation des Clusters (Louvain)	19
Figure 11 : Comparaison entre K-Means et Louvain	19

Table des matières

Liste des figures	2
Table des matières.....	3
Introduction	1
I. Réseaux Complexes :	2
I.1 Définition	2
I.2 Exemples pertinents	2
I.3 Réseaux routiers	2
I.4 Réseaux biologiques	2
I.5 Réseaux sociaux.....	2
I.6 Caractéristiques essentielles.....	3
II. Détection de Communauté :	4
II.1 Définition :	4
II.2 Les mesures de centralité :	4
II.3 Centralité de Degré :	4
II.4 Centralité de Proximité :	4
II.5 Centralité d'Intermédiation	4
II.6 Centralité de Vecteur Propre.....	5
II.7 Centralité et Détection de Communautés.....	5
III. Étude de L'Existence :	6
III.1 Les Travaux les Plus Récents :	6
III.2 Défis Actuels et Perspectives :	6
IV. Algorithmes Utilisés	8
IV.1 Algorithme de K-Means :	8
IV.1.1 Définition :	8
IV.1.2 Principe :	8
IV.2 Algorithme Louvain :	9
IV.2.1 Définition	9
IV.3.2 Principe :	9
V. Réalisation du projet	10
V.1 Outils et bibliothèques	10
V.1.1 Bibliothèques :	10
V.1.2 Outils :	11
V.2 Réseau Global du Dataset Facebook :	11
V.2.1 <i>Résultat avec algorithme K-Means</i> :	12
V.2.2 <i>Résultat avec algorithme Louvain</i> :	14
V.3 Sous-réseau du Dataset Facebook :	15
V.3.1 <i>Résultat avec algorithme K-Means</i> :	16
V.3.2 <i>Résultat avec algorithme Louvain</i> :	17
V.3.3 <i>Comparaison</i> :	19
VI. Conclusion	21
VII. Référence :	22

Introduction

Dans ce monde interconnecté, la compréhension des réseaux complexes s'avère essentielle. Ces réseaux, omniprésents dans notre quotidien, de la diffusion des informations sur les réseaux sociaux à la complexité des interactions biologiques, offrent un aperçu fascinant sur la manière dont les éléments sont interconnectés. L'un des aspects les plus intéressants de l'analyse de ces réseaux est la détection de communautés : une technique permettant d'identifier des groupes de nœuds présentant des connexions plus denses en interne qu'avec l'extérieur du groupe. Cette méthode révèle non seulement les sous-structures cachées mais aussi les tendances et les dynamiques au sein des réseaux.

Notre projet se concentre sur l'application de divers algorithmes de détection de communautés sur des ensembles de données choisis d'un réseau social, permettant ainsi de dévoiler les multiples aspects de ce réseau complexe. En combinant analyse théorique et pratique, ce rapport vise à enrichir notre compréhension des réseaux et à mettre en lumière les potentiels des différentes approches algorithmiques.

Pour ce faire, nous avons organisé notre rapport en quatre parties, chacune ciblant un aspect clé de notre étude, allant de l'examen des algorithmes à l'interprétation des résultats, en passant par la méthodologie et l'analyse des données.

I. Réseaux Complexes :

I.1 Définition

Un réseau complexe est un système caractérisé par des interconnexions irrégulières et souvent dynamiques entre ses composants. Contrairement aux réseaux simples aux connexions prévisibles, les réseaux complexes affichent des structures variées et des interactions diversifiées. Ils reflètent la complexité des systèmes réels, comme les réseaux sociaux et ils sont essentiels pour comprendre des phénomènes tels que la diffusion de l'information ou la dynamique des écosystèmes.

I.2 Exemples pertinents

Parmi les réseaux complexes les plus connus, on trouve les réseaux routiers, biologiques et sociaux, chacun illustrant des aspects uniques de ces structures interconnectées.

I.3 Réseaux routiers

Ces réseaux représentent le système de routes et de voies de circulation qui connectent différents lieux géographiques. Ils sont caractérisés par leur structure en maillage, où les intersections et les autoroutes fonctionnent comme des nœuds et des arêtes, respectivement. La complexité de ces réseaux réside dans leur capacité à modéliser la circulation, les flux de trafic et les itinéraires optimaux.

I.4 Réseaux biologiques

Ils comprennent les réseaux neuronaux, où les neurones sont interconnectés par des synapses, et les réseaux génétiques, où les gènes interagissent de diverses manières. Ces réseaux sont essentiels pour comprendre les processus biologiques complexes, comme le fonctionnement du cerveau ou les mécanismes de régulation génétique.

I.5 Réseaux sociaux

Sites comme Facebook ou Twitter illustrent ce type de réseau, où les individus (nœuds) interagissent via des relations sociales (arêtes). Ces réseaux reflètent la complexité des interactions humaines et sont utilisés pour étudier des phénomènes tels que la diffusion de l'information, la formation de groupes sociaux, ou même la propagation de tendances et d'idées.

I.6 Caractéristiques essentielles

Les réseaux complexes sont caractérisés par des traits distincts qui influencent leur structure et leur fonctionnement

- **Distribution des Degrés** : Typique des réseaux complexes, la distribution des degrés suit souvent une loi de puissance. Cela signifie que quelques nœuds ont un grand nombre de connexions, tandis que la majorité des nœuds en ont relativement peu. Cette distribution inégale joue un rôle crucial dans la façon dont l'information ou l'influence se propage à travers le réseau.
- **Formation de Clusters** : Les réseaux complexes ont une tendance marquée à la formation de clusters ou de communautés. Dans ces groupes, les nœuds sont plus densément connectés entre eux qu'avec le reste du réseau. Ce phénomène reflète la nature des interactions au sein du réseau et est essentiel pour comprendre sa structure locale et globale.

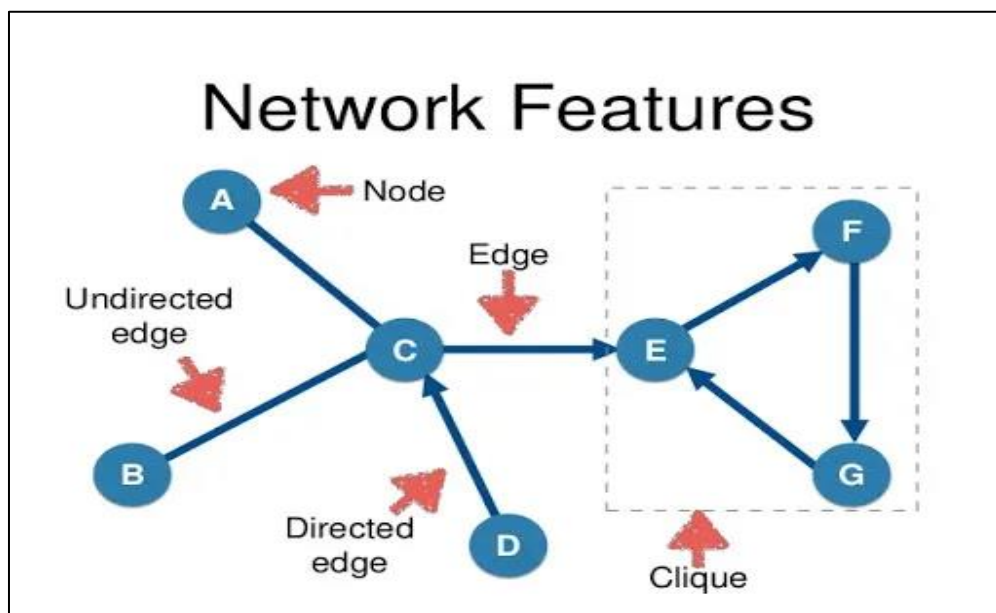


Figure 1 : Caractéristiques des réseaux

II. Détection de Communauté :

II.1 Définition :

La détection de communauté dans les réseaux complexes est une méthode analytique visant à regrouper les nœuds en fonction de la similarité de leurs connexions. Une 'communauté' dans ce contexte est définie comme un ensemble de nœuds qui sont plus densément connectés entre eux qu'avec le reste du réseau. L'objectif est de découvrir la structure modulaire du réseau, qui peut souvent révéler des informations sur la fonctionnalité, les interactions sociales, ou les flux d'information au sein du réseau. Les communautés peuvent être comprises comme des unités fonctionnelles au sein d'un réseau plus vaste, chacune possédant ses propres caractéristiques et son propre rôle.

II.2 Les mesures de centralité :

Les mesures de centralité sont des indicateurs clés utilisés pour identifier les nœuds les plus importants au sein d'un réseau complexe. Elles permettent d'évaluer l'influence ou la prééminence d'un nœud dans un réseau en fonction de sa position et de ses connexions. Voici les principales mesures de centralité.

II.3 Centralité de Degré :

Cette mesure se concentre sur le nombre de connexions directes qu'un nœud a avec d'autres dans le réseau. Un nœud avec un degré élevé peut être considéré comme un 'hub' ou un point central, jouant un rôle crucial dans la connectivité du réseau.

II.4 Centralité de Proximité :

Elle examine à quel point un nœud est proche de tous les autres nœuds du réseau en calculant la somme des distances les plus courtes. Les nœuds avec une centralité de proximité élevée peuvent rapidement interagir avec tous les autres nœuds et sont souvent stratégiques pour optimiser la communication au sein du réseau.

II.5 Centralité d'Intermédiation :

Cette mesure identifie les nœuds qui servent de ponts ou d'intermédiaires le long des chemins les plus courts entre les paires de nœuds. Les nœuds ayant une forte centralité d'intermédiation contrôlent souvent le flux d'informations ou de ressources et peuvent être critiques pour la cohésion du réseau.

II.6 Centralité de Vecteur Propre

En prenant en compte non seulement le nombre de connexions qu'un nœud a mais aussi la centralité des nœuds auxquels il est connecté, cette mesure attribue une importance plus grande aux connexions avec d'autres nœuds influents. Elle est souvent utilisée pour identifier les nœuds influents dans les réseaux de communication et de recommandation.

II.7 Centralité et Détection de Communautés

La relation entre les mesures de centralité et la détection de communauté est essentielle pour une analyse approfondie des réseaux complexes. Cette section explore comment les caractéristiques des nœuds influencent la formation et la dynamique des communautés.

- **Impact sur la Formation de Communautés :** Les nœuds avec une haute centralité de degré tendent à être des points autour desquels les communautés se forment. Ces derniers facilitent une densité de connexion élevée au sein d'une communauté, influençant ainsi sa structure et sa cohésion.
- **Rôle dans la Dynamique des Communautés :** Les nœuds ayant une forte centralité d'intermédierité sont souvent des points de passage clés entre différentes communautés. Ils jouent un rôle crucial dans la facilitation de la communication et de l'échange d'informations entre groupes distincts, ce qui peut avoir des implications significatives pour la compréhension de la structure globale du réseau.
- **Utilisation Stratégique dans l'Analyse de Réseau :** En détection de communauté, utiliser stratégiquement les mesures de centralité peut conduire à une meilleure compréhension des réseaux. Par exemple, en identifiant et en analysant des nœuds avec une centralité élevée, les chercheurs peuvent déterminer des points clés qui influencent fortement la structure et la dynamique du réseau, permettant ainsi une analyse plus ciblée et efficace.

Cette interaction entre les mesures de centralité et la détection de communauté révèle des aspects cruciaux de la structure et du fonctionnement des réseaux, offrant ainsi une perspective plus riche pour l'analyse des réseaux complexes.

III. Étude de L'Existence :

La détection des communautés dans les réseaux sociaux constitue un domaine de recherche dynamique, évoluant en réponse aux avancées technologiques et aux nouveaux défis posés par les réseaux sociaux contemporains. Cette section examine les tendances actuelles ainsi que les défis persistants dans ce domaine.

III.1 Les Travaux les Plus Récents :

- **Constructions sociales sur le covid-19 au Sénégal 2024 : logiques de communication net appropriation des discours - ID DIONE ;** La détection des communautés émerge comme un outil pour décrypter les logiques de communication et l'appropriation des discours liés au COVID-19 au Sénégal. Elle permet de mettre en lumière les dynamiques sociales à l'œuvre sur les réseaux sociaux, où la désinformation et la mésinformation peuvent influencer la perception collective de la crise.
- **De l'information aux industries culturelles 2024 : l'hypothèse chahutée de la bulle de filtre - J Farchy, S Tallec ;** La détection des communautés offre un regard privilégié sur la formation de bulles de filtre dans les réseaux sociaux. Elle devient ainsi un instrument indispensable pour comprendre comment l'information circule et est filtrée, affectant la manière dont les industries culturelles sont perçues et consommées.
- **La violence numérique en milieu scolaire 2024 : Un défi croissant à relever - B Iman ;** L'importance de la détection des communautés est soulignée dans la compréhension et la prévention de la violence numérique en milieu scolaire. Cet outil devient crucial pour analyser les dynamiques sociales en ligne et élaborer des stratégies efficaces face à ce défi croissant.

III.2 Défis Actuels et Perspectives :

- **Taille et Complexité des Réseaux :** La croissance exponentielle des réseaux sociaux pose un défi majeur, car la taille et la complexité de ces réseaux rendent difficile la détection précise des communautés. Les méthodes existantes peuvent être limitées dans leur capacité à gérer ces réseaux massifs.

- **Détection de Communautés Chevauchantes :** La présence d'utilisateurs appartenant à plusieurs communautés représente un défi significatif. Les approches traditionnelles peuvent avoir du mal à traiter ces situations de chevauchement, nécessitant des méthodes plus sophistiquées.
- **Protection de la Vie Privée :** La protection de la vie privée des utilisateurs représente un défi éthique majeur. Les chercheurs doivent élaborer des méthodes qui garantissent la confidentialité tout en maintenant la qualité des résultats.

IV. Algorithmes Utilisés

IV.1 Algorithme de K-Means :

IV.1.1 Définition :

L'algorithme K-Means est une technique de clustering utilisée pour regrouper des nœuds similaires dans un réseau social en K communautés distinctes. K-Means repose sur la partition du réseau en K clusters en minimisant la somme des carrés des distances intra-cluster.

IV.1.2 Principe :

1. **Initialisation des Centres de Gravité :** On commence par choisir aléatoirement K nœuds comme centres initiaux des clusters.
2. **Attribution des Nœuds aux Clusters les Plus Proches :** Chaque nœud est attribué au cluster dont le centre est le plus proche en termes de distance euclidienne.
3. **Recalcul des Centres de Gravité :** Les centres des clusters sont recalculés comme la moyenne des coordonnées de tous les nœuds appartenant à chaque cluster.
4. **Répétition des Étapes 2-3 :** Les étapes d'attribution et de recalcul sont répétées jusqu'à convergence, où les centres des clusters ne changent plus de manière significative.
5. **Optimisation de la Partition :** L'algorithme vise à minimiser la somme des carrés des distances intra-cluster, optimisant ainsi la partition du réseau en K clusters.

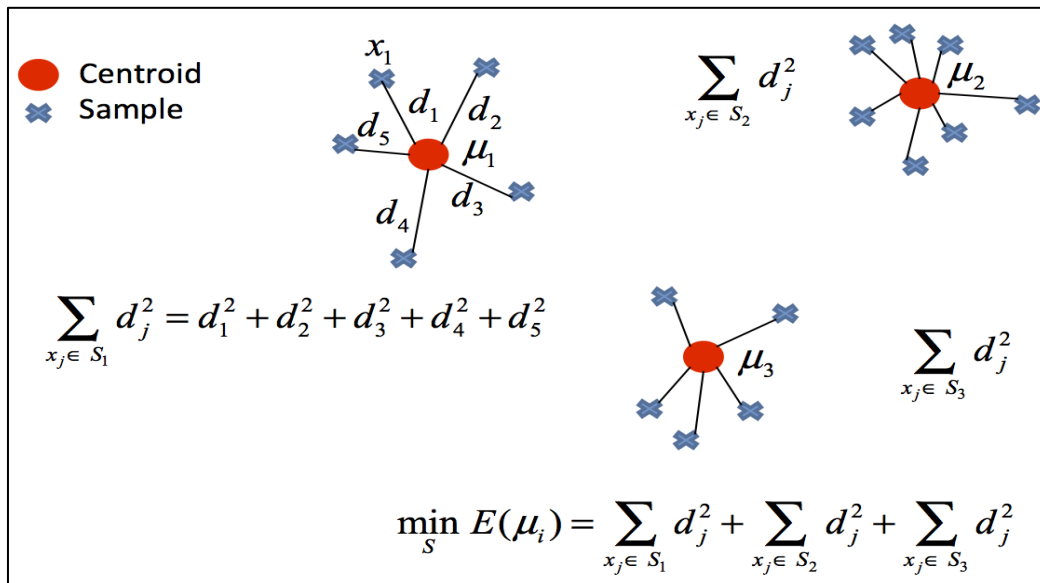


Figure 2 : K-Means Principe

IV.2 Algorithme Louvain :

IV.2.1 Définition :

L'algorithme de Louvain est une méthode de détection de communauté basée sur la maximisation de la modularité. Il est bien connu pour son efficacité et sa capacité à traiter de grands réseaux.

IV.3.2 Principe :

1. **Choix d'une Partition Initiale Aléatoire :** On commence par attribuer aléatoirement chaque nœud à une communauté.
2. **Itération sur les Nœuds :** On itère sur chaque nœud du réseau, en examinant individuellement les possibilités de réaffectation communautaire.
3. **Calcul du Changement de Modularité :** Pour chaque nœud, on calcule le changement de modularité qui résulterait de son déplacement vers une autre communauté. Cette étape évalue l'impact potentiel de différentes réaffectations sur la qualité globale de la structure communautaire.
4. **Optimisation de la Modularité :** On choisit le déplacement qui apporte la plus forte augmentation de la modularité. Ce processus est fondé sur la recherche du meilleur gain en termes de densité de connexions internes à la communauté versus les connexions externes.
5. **Répétition Jusqu'à Optimisation :** Les étapes 2 à 4 sont répétées de façon itérative jusqu'à ce qu'aucune amélioration supplémentaire de la modularité ne soit possible.

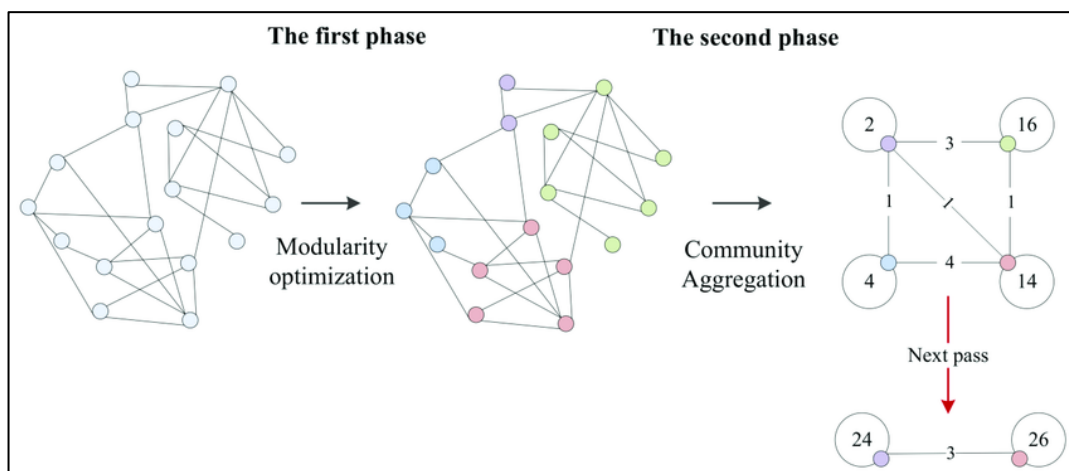


Figure 3 : Louvain Principe

V. Réalisation du projet

Dans cette section, nous présentons et comparons les résultats obtenus avec les algorithmes de K-Means et Louvain appliqués sur Facebook Dataset. Notre analyse se focalise sur la modularité des communautés détectées, offrant un aperçu de l'efficacité de chaque algorithme.

V.1 Outils et bibliothèques

V.1.1 Bibliothèques :

- **Pandas**

Pandas est une bibliothèque open-source en Python spécialisée dans la manipulation et l'analyse de données tabulaires. Elle offre des structures de données flexibles comme les DataFrame et des outils performants pour nettoyer, transformer et analyser des données.

- **NetworkX**

NetworkX est une bibliothèque open-source en Python dédiée à la création, la manipulation et l'analyse de réseaux complexes. Elle offre des outils pour représenter des graphes et effectuer diverses opérations telles que la recherche de chemins, l'identification de communautés, et l'analyse de la topologie du réseau.

- **Matplotlib**

Matplotlib est une bibliothèque en Python utilisée pour créer des visualisations graphiques, telles que des graphiques, des diagrammes en barres, des histogrammes, des nuages de points, etc. Elle offre un large éventail de fonctionnalités permettant aux utilisateurs de personnaliser et de contrôler divers aspects de leurs visualisations.

- **Scikit-learn**

Scikit-learn est une bibliothèque open-source en Python dédiée à l'apprentissage automatique. Elle propose une interface simple et efficace pour l'entraînement de modèles, l'évaluation de leur performance, ainsi que la réalisation de diverses tâches liées à l'analyse de données et à la construction de modèles prédictifs.

V.1.2 Outils :

- **Jupyter**

Jupyter est une plateforme open-source qui permet la création et le partage de documents interactifs intégrant du code, des équations, des visualisations et du texte. Elle prend en charge divers langages de programmation, favorisant l'exploration interactive des données et le développement de code dans des cellules.

- **Gephi**

Gephi est une application open-source spécialisée dans la visualisation et l'exploration de réseaux complexes. Elle offre des outils graphiques avancés permettant d'analyser la structure et les relations au sein de réseaux, facilitant ainsi la compréhension visuelle des données interconnectées.

- **Ananconda**

Anaconda est une distribution Python open-source et une plateforme de gestion de paquets, simplifiant l'installation et la gestion d'outils scientifiques et analytiques. Intégrant des packages populaires et un gestionnaire de paquets, elle offre un environnement complet pour le développement et la gestion de projets en science des données et en apprentissage automatique.

V.2 Réseau Global du Dataset Facebook :

Dans cette étude, nous avons utilisé Facebook Dataset. Facebook Dataset a été sélectionné pour son authenticité et sa pertinence dans la représentation des relations sociales en ligne. Ces données fournissent un aperçu unique des liens d'amitié sur la plateforme, permettant ainsi d'explorer les structures communautaires au sein du réseau.

Composé de 4039 nœuds et 88234 arêtes, ce réseau représente des relations sociales plus précisément en Facebook. Les données étaient déjà formatées pour une analyse directe, donc on n'a pas passé par l'étape de préparation de données.

Passant maintenant à la visualisation du réseau complexe de cette dataset :

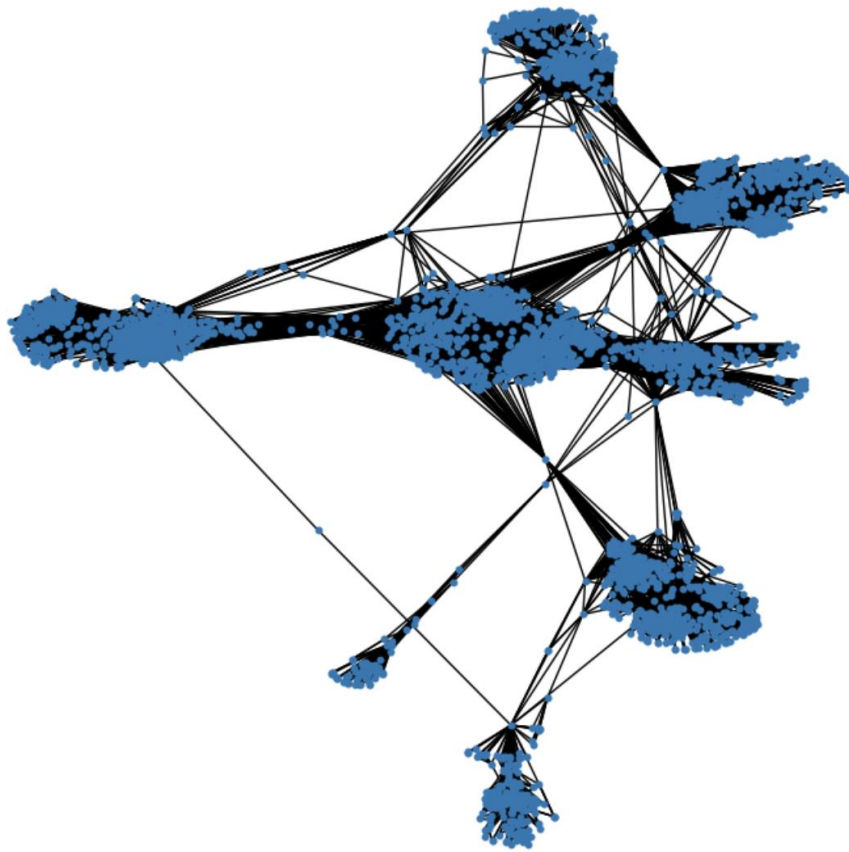


Figure 4 : Réseau du Facebook Dataset

Dans cette partie essentielle de notre étude, nous appliquons les algorithmes de Louvain et K-Means sur le dataset Facebook, en mettant en lumière les résultats obtenus et les particularités de chaque algorithme.

V.2.1 Résultat avec algorithme K-Means :

- Nous avons extrait des caractéristiques cruciales des nœuds du réseau, comme la centralité degré, d'intermédiarité et de proximité. Le nœud le plus influent a été identifié grâce à la centralité d'intermédiarité. Ces données ont été structurées dans un DataFrame et normalisées pour une analyse cohérente.
- En utilisant l'algorithme K-means avec dis clusters, nous avons regroupé les nœuds similaires. La visualisation du graphe met en évidence ces clusters par des couleurs distinctes, avec le nœud le plus influent en jaune.

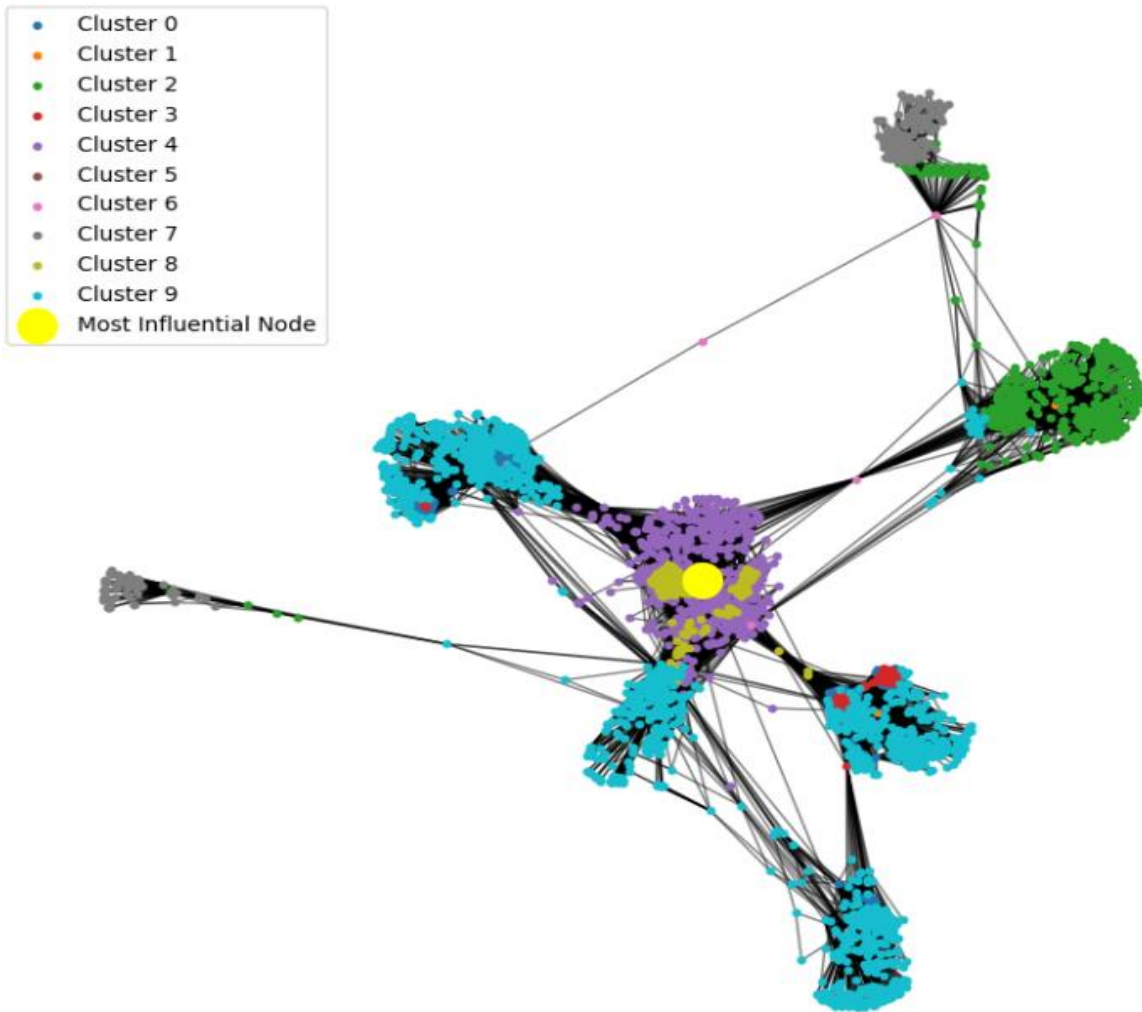


Figure 5 : Réseau avec Visualisation des Clusters (K-Means)

V.2.2 Résultat avec algorithme Louvain :

- L'algorithme de Louvain a été appliqué pour détecter des clusters dans le réseau. Chaque nœud a été attribué à un cluster spécifique, formant ainsi des regroupements distincts. Le nœud le plus influent selon l'algorithme Louvain a été identifié.
- En total, 16 clusters ont été détectés dans le réseau. La visualisation graphique met en évidence ces clusters avec des couleurs distinctes, et le nœud le plus influent est mis en évidence en jaune.

Ci-dessous, nous présentons une visualisation des clusters détectées dans le réseau à l'aide de l'algorithme de Louvain, illustrant les différentes répartitions.

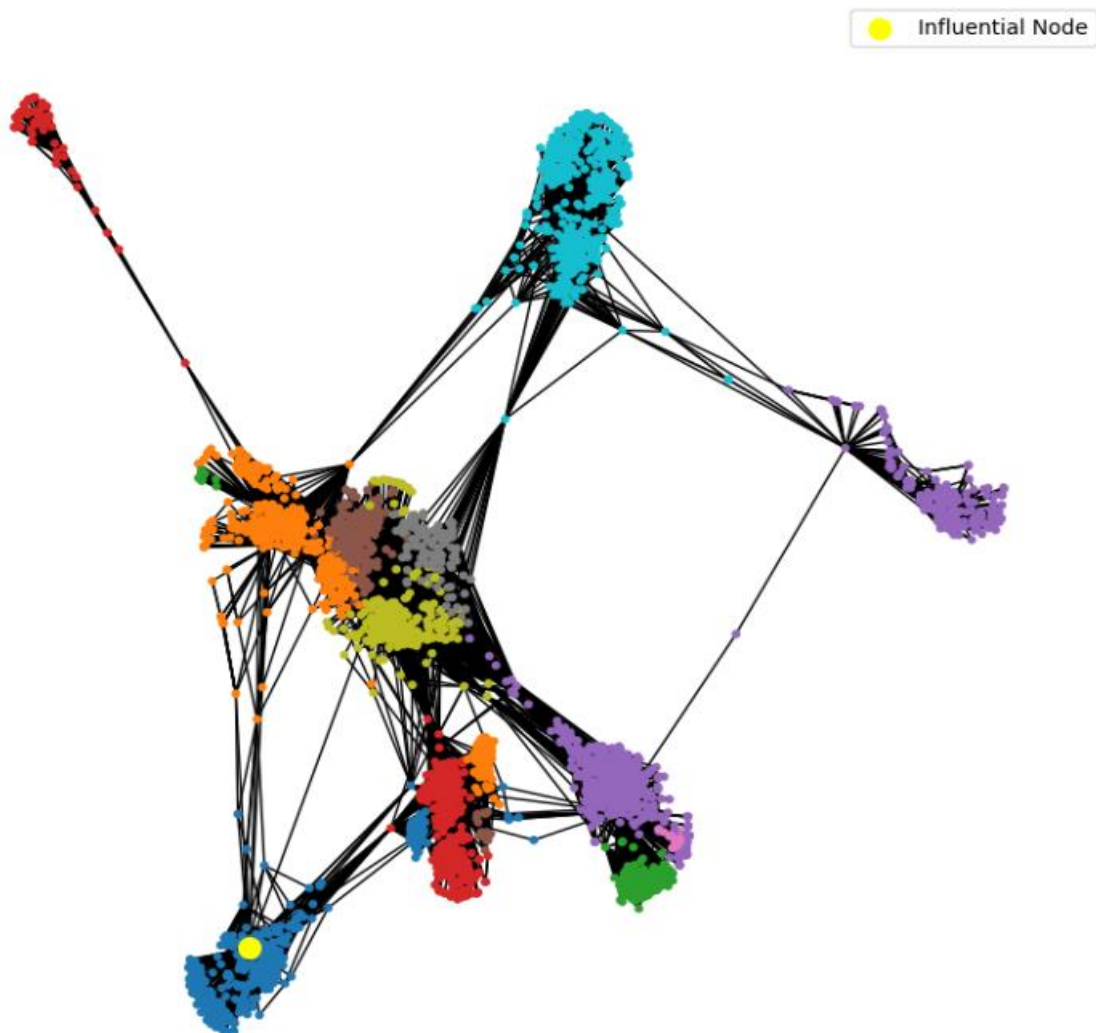


Figure 6 : Réseau avec Visualisation des Clusters (Louvain)

V.3 Sous-réseau du Dataset Facebook :

- Dans le but de simplifier la visualisation et de mieux comprendre la structure du réseau, un sous-graphe représentatif a été créé en choisissant aléatoirement 200 nœuds parmi l'ensemble du réseau initial. Cette approche permet de focaliser l'analyse sur un échantillon significatif du réseau complet, facilitant ainsi l'identification de clusters et de relations plus claires.
- Le sous-graphe généré compte 200 nœuds et 962 arêtes, offrant une vue condensée mais représentative du réseau.

Passant maintenant à la visualisation du sous-réseau complexe de cette dataset en utilisant la bibliothèque Matplotlib :

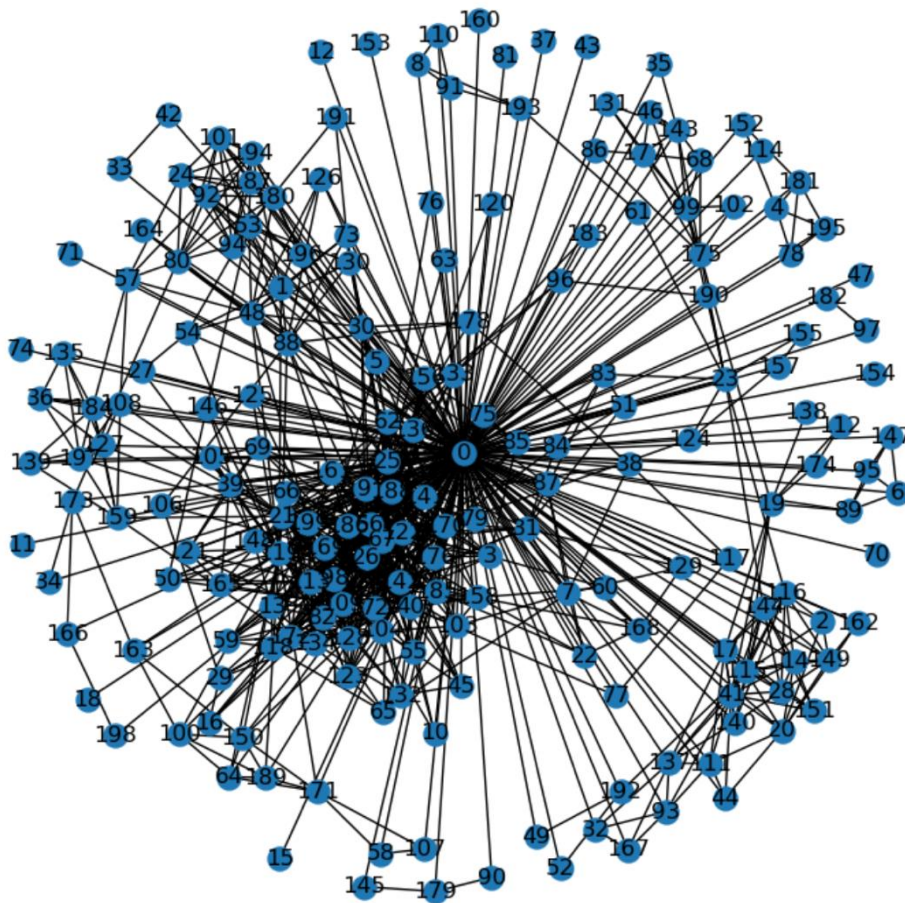


Figure 7 : Sous-Graphe Représentatif du Dataset Facebook

Pour une visualisation plus détaillée et expressive du Sous-Graphe Représentatif, une représentation graphique a été générée avec l'outil Gephi. La figure ci-dessous offre une perspective plus claire sur la structure du réseau.

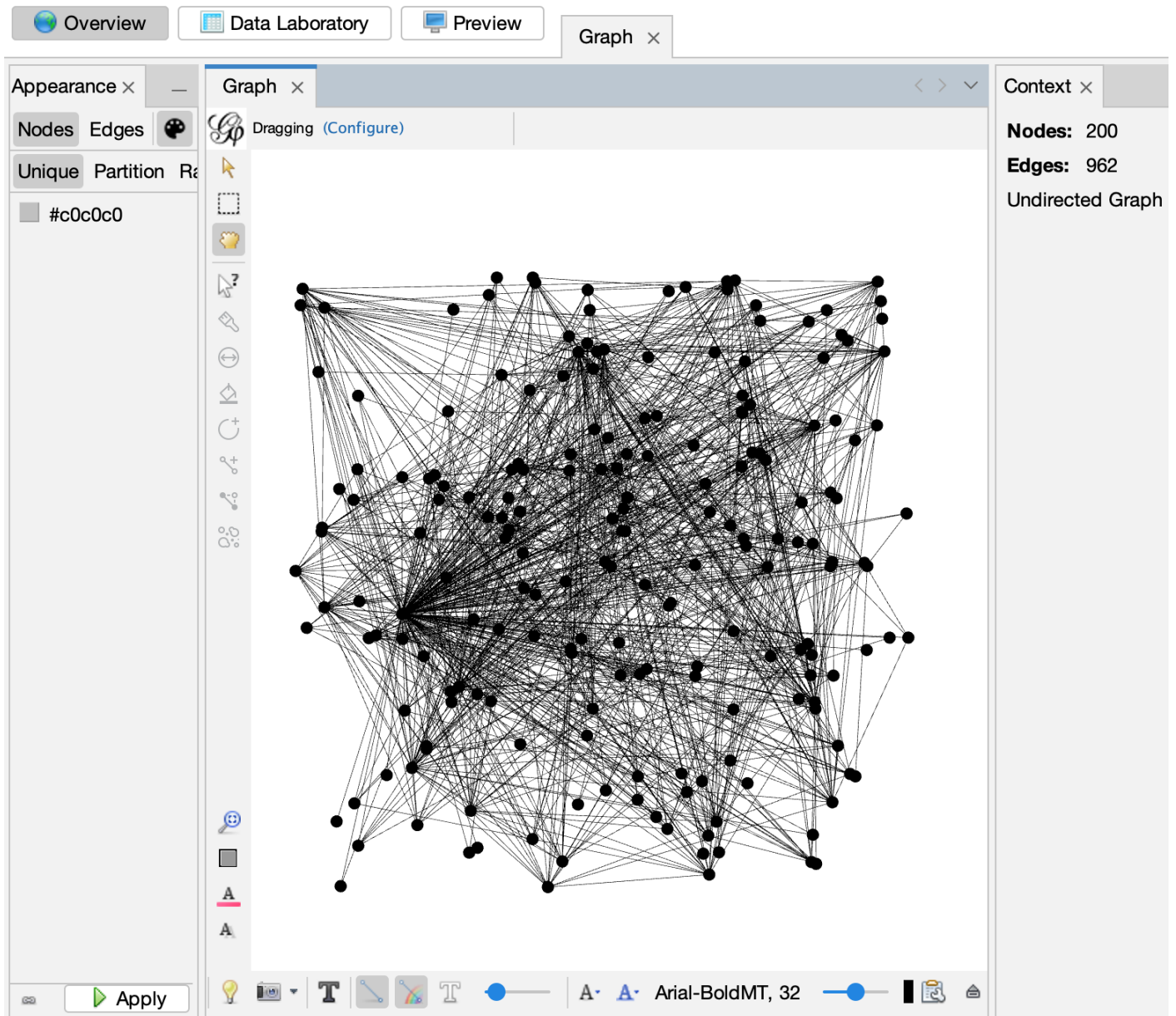


Figure 8 : Visualisation du Sous-Graphe Représentatif avec Gephi

V.3.1 Résultat avec algorithme K-Means :

Afin d'analyser plus en détail le Sous-Graphe Représentatif, l'algorithme K-Means a été appliqué aux caractéristiques des nœuds normalisées. Avec une détermination de 4 clusters, les

nœuds ont été regroupés en fonction de leurs similarités, permettant une meilleure compréhension des structures sous-jacentes dans le sous-réseau.

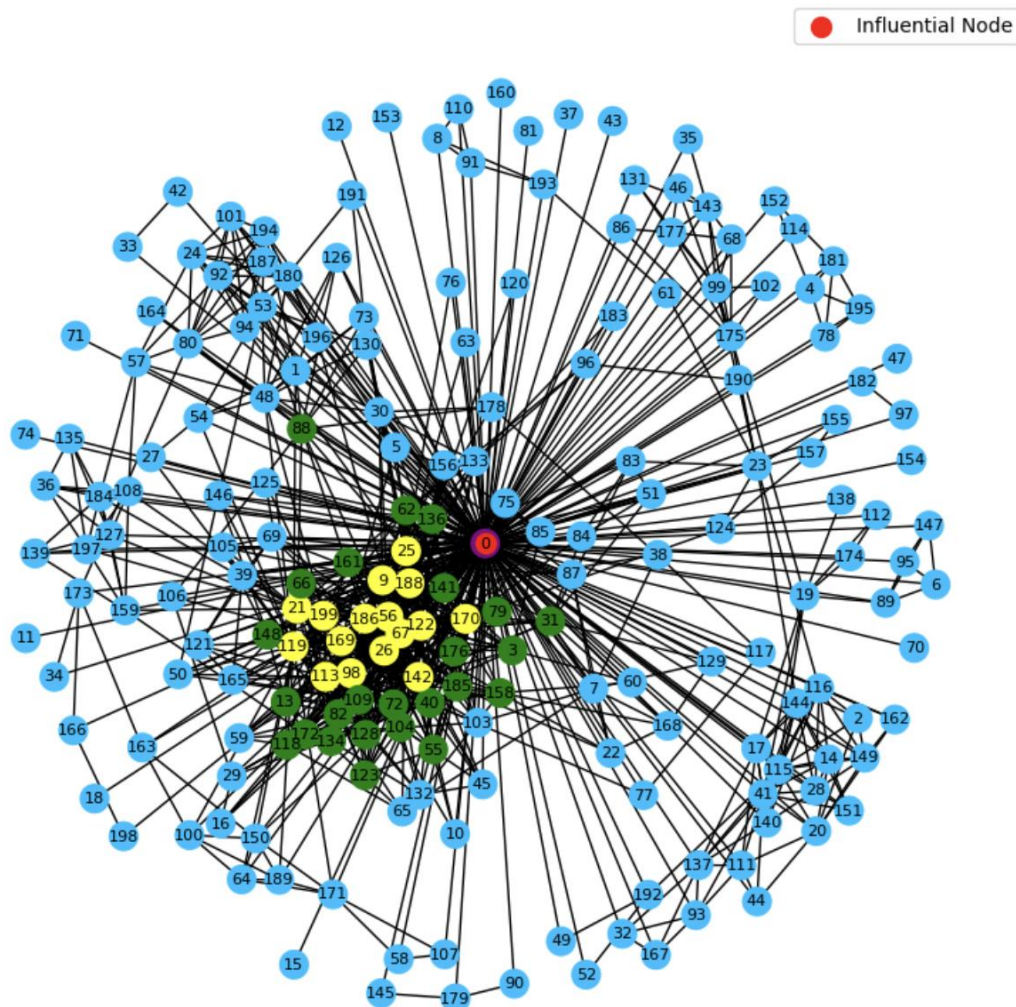


Figure 9 : Sous-Réseau avec Visualisation des Clusters (K-Means)

V.3.2 Résultat avec algorithme Louvain :

Afin d'explorer davantage la structure du Sous-Graphe Représentatif, l'algorithme de Louvain a été appliqué pour identifier des clusters significatifs parmi les nœuds. La visualisation graphique ci-dessous présente ces clusters détectés par Louvain, avec une mise en évidence particulière du nœud le plus influent marqué en jaune. Le Sous-Graphe Représentatif semble

avoir été regroupé en 7 clusters distincts selon l'algorithme Louvain, apportant des informations supplémentaires sur la structure sous-jacente du réseau.

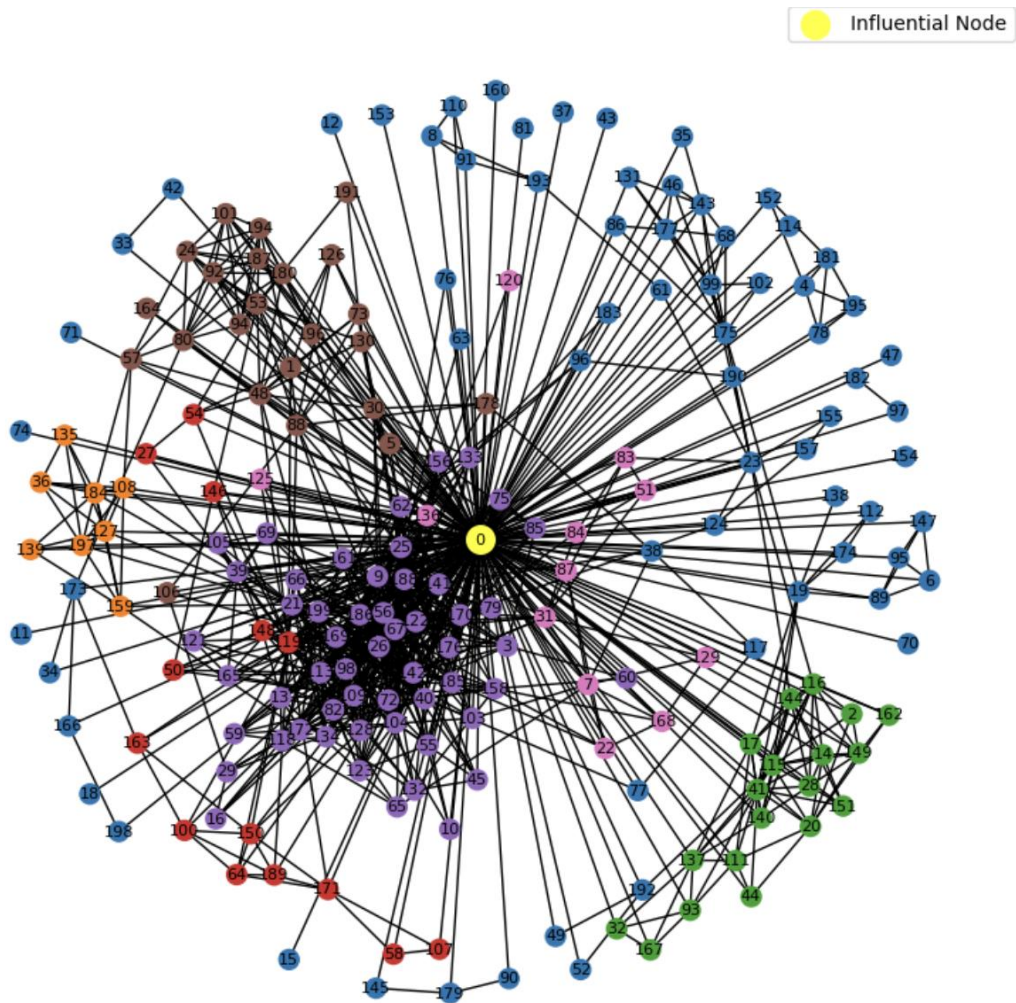


Figure 10 : Sous-Réseau avec Visualisation des Clusters (Louvain)

V.3.3 Comparaison :

```
silhouette_kmeans = silhouette_score(scaled_features, clusters)
print(f"Le coefficient de silhouette pour KMeans est : {silhouette_kmeans}")

# Calculer le coefficient de silhouette pour Louvain
silhouette_louvain = silhouette_score(scaled_features, node_colors_louvain)
print(f"Le coefficient de silhouette pour Louvain est : {silhouette_louvain}")
```

Le coefficient de silhouette pour KMeans est : 0.6756361435266863
Le coefficient de silhouette pour Louvain est : -0.17071139869784868

```
# Calcul de la modularité pour KMeans
modularity_kmeans = community.modularity(kmeans_cluster_dict, subgraph)
print(f"La modularité pour KMeans est : {modularity_kmeans}")
modularity_louvain = community.modularity(louvain_partition, subgraph)
print(f"La modularité pour Louvain est : {modularity_louvain}")
```

La modularité pour KMeans est : 0.11828322405245485
La modularité pour Louvain est : 0.45823788365368406

Figure 11 : Comparaison entre K-Means et Louvain

Modularité :

- KMeans : 0.118
- Louvain : 0.458

La modularité évalue la qualité de la division en clusters. Pour notre sous-graphe, la modularité est plus élevée avec Louvain (0.458), suggérant une meilleure séparation en clusters par rapport à KMeans (0.118). Louvain semble avoir réussi à identifier des communautés plus distinctes dans notre graphe.

Coefficient de Silhouette :

- KMeans : 0.676
- Louvain : -0.17

Le coefficient de silhouette mesure à quel point les nœuds dans le même cluster sont similaires par rapport à ceux des autres clusters. KMeans affiche un coefficient de silhouette plus élevé (0.676), indiquant que les nœuds au sein des clusters sont plus similaires. Cependant, le coefficient de silhouette pour Louvain est négatif (-0.17), suggérant que la structure détectée peut ne pas être aussi claire.

Conclusion:

- La modularité suggère que Louvain excelle dans la détection de communautés distinctes.
 - Le coefficient de silhouette indique que, bien que KMeans ait des clusters plus homogènes, Louvain peut présenter des clusters moins bien séparés.
- ✓ si l'objectif principal est la détection de communautés, Louvain semble plus performant en termes de modularité. Cependant, pour l'homogénéité intra-cluster, KMeans a un avantage en termes de coefficient de silhouette. Considérer ces deux mesures ensemble offre une perspective complète de la performance des algorithmes dans notre contexte d'analyse de réseau.

VI. Conclusion

Au terme de cette étude approfondie sur le réseau du Dataset Facebook, l'application des algorithmes KMeans et Louvain a fourni des perspectives cruciales sur la structuration de ce réseau social. Les résultats obtenus révèlent des nuances intéressantes dans la détection de clusters. Alors que Louvain a démontré une meilleure capacité à identifier des communautés distinctes, KMeans a mis en évidence des clusters plus homogènes en termes de caractéristiques des nœuds.

Cette analyse met en lumière l'importance de choisir des méthodes de clustering en fonction des objectifs spécifiques de l'analyse du réseau. Les métriques telles que la modularité et le coefficient de silhouette ont été cruciales pour évaluer la performance des algorithmes dans ce contexte.

En conclusion, cette étude contribue significativement à notre compréhension des structures sous-jacentes du réseau Facebook, mettant en évidence les avantages respectifs de Louvain et KMeans dans ce contexte spécifique. Ces découvertes ouvrent des perspectives pour des recherches futures visant à explorer davantage les mécanismes complexes des réseaux sociaux.

VII. Référence :

Google Scholar: <https://scholar.google.com/scholar>

K-Means: <https://scikitlearn.org/stable/modules/generated/sklearn.cluster.KMeans.html>

Louvain : <https://louvain-igraph.readthedocs.io>

Networkx: <https://networkx.org/documentation>

Gephi: <https://docs.gephi.org>