Faculté
des Sciences
Department of Informatics

UMONS
Université de Mons

Les cliques maximales dans les graphes dispersés

Lecture et rédaction scientifiques

Mémoire réalisé par Bal Sébastien
pour l'obtention du diplôme de Master en Informatics

Service : Master Sciences Informatique
Directeur : Mr Mélot

Année académique 2023-2024

Faculté des Sciences ◆ Université de Mons ◆ Place du Parc 20 ◆
B-7000 Mons

Table des matières

Introduction

Dans cette partie, une brève mise en bouche du sujet avec une partie qui présentera les différents mots de vocabulaire que l'on utilisera dans l'article. Ensuite, une mise en avant du fonctionnement de l'algorithme ainsi qu'une démonstration de différents schémas et de l'utilisation des données pour montrer l'efficacité de cet algorithme. Une explication sera fournie avec des commentaires sur l'algorithme qui sera intégré avec python ainsi que les graphes qui auront été générés.

Les bases et notions de vocabulaire

Graphes

Un graphe est composé d'un ensemble de points, appelés sommets, ceux-ci sont reliés entre eux par des lignes que l'on appelle des arêtes. Un sommet est représenté par un point et une arête par une ligne.

On donne généralement comme définition pour un graphe cette notation : $G = (V, E)$. G est un couple (V, E) dans lequel, V est un ensemble fini de sommets et E est un ensemble d'arêtes, où chaque arête est un sous-ensemble de sommets de V noté : $\{v_i, v_j\} \in V^2$.

Une arête est une entité caractérisée par une paire de sommets $\{v_i, v_j\}$. Ces deux sommets sont considérés comme adjacents. L'ensemble de ces sommets adjacents $v_i \in V$ est représenté sous la notation $Adj(v_i) = \{v_i \in V, \{v_i, v_j\} \in E\}$

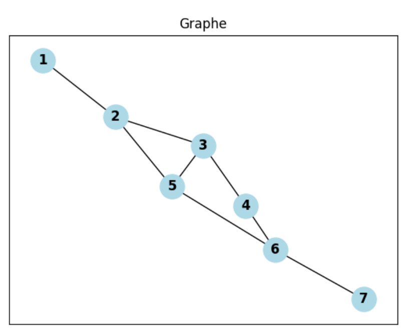

Figure 1 – Représentation d'un graphe

Une Clique

Une clique, au point de vue de la théorie des graphes, représente un sous-ensemble de sommets dans lequel chaque paire de sommets est reliée[2].*La taille d'une clique* est un ensemble de k nœuds dans lequel chaque paire de nœuds est reliée par une arête. Ceux-ci sont tous connectés, on parle de *sous-graphe* complet.

Un sous-graphe est un graphe se trouvant dans un autre graphe, la figure 2 représente un sous-graphe en vert dans un graphe en bleu.

Sur la figure 2, on peut constater un graphe qui possède 7 sommets, les sommets en vert, numéroté 2,3,5, forment une clique de taille 3. Plusieurs méthodes existent pour détecter des cliques dans un graphe donné, celles-ci sont généralement complexes à utiliser. L'application la plus importante dans la détection de cliques est de trouver le nombre maximum de cliques dans un graphe.

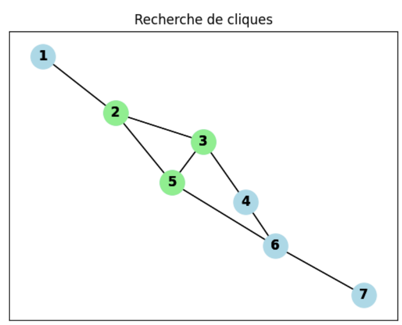

FIGURE 2 – Représentation d'une Clique

Dégénérescence

La dégénérescence dans la théorie des graphes, est un paramètre. Celle-ci est la plus petite valeur (k) de façon qu'il soit k-dégénéré, sous-graphe contenant au moins un sommet de *degré* supérieur à k[1]. Comme le défini Wikipédia[3], *un degré* d'un sommet dans la théorie des graphes, c'est le nombre de liens (arêtes) reliant ce sommet.

L'article de Seidman\cite(seidman), présente qu'un k-core doit avoir au moins $k+1$ points, ainsi que les points des différents k-core ne peuvent pas être adjacents. Prenons les exemples qui suivent, en changeant la valeur de k, on obtient différent sous-graphe.

Le premier graphe est celui avec $k = 0$, cela reprend tous les points du graphe à la figure 3.

Le graphe de la figure 4 illustre un graphe de dégénérescence avec $k = 1$. Tous les sommets sont inclus dans l'ensemble de ce sous-graphe, étant donné que

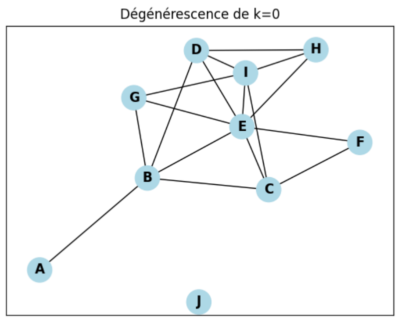

FIGURE 3 – Dégénérescence 0

chaque sommet est adjacent à au moins un autre sommet. La non-conformité à la condition $k = 1$ surviendrait si un sommet n'était pas relié à un autre. Par exemple, le sommet J ne respecte pas cette condition, car il est isolé, n'ayant aucune arête incidente avec d'autres sommets.

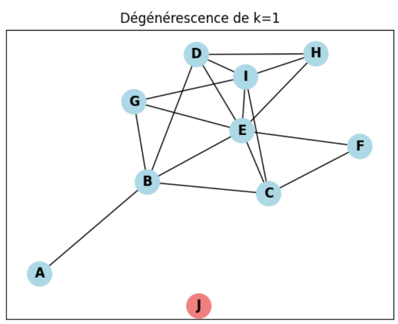

FIGURE 4 – Dégénérescence 1

Dans le graphe de la figure 5, avec $k = 2$, nous notons que les sommets A et J ne sont pas inclus dans le sous-graphe. Ces sommets sont affichés en rouge pour indiquer qu'ils ne satisfont pas la contrainte de degré minimal égal ou supérieur à 2, conforme à la définition d'un 2-core.

Pour une dégénérescence dans laquelle $k = 3$, le graphe à la figure 6, ne regroupe pas dans le sous-graphe les sommets A, F et J. Pour cette dernière représentation, c'est le point F qui ne rentre pas dans les bonnes conditions.

La valeur maximum qui a été trouvée pour le graphe à la figure 6, est la valeur de $k = 3$, si $k = 4$ alors tous les sommets ne feraient plus partie du sous-graphe, car aucun ne respectera la condition minimum de 3 arêtes.

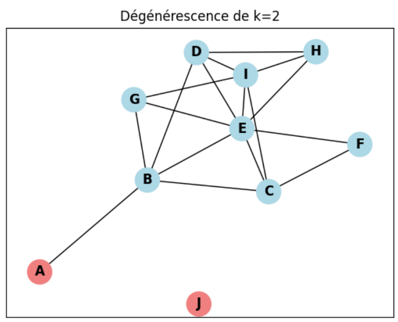

Figure 5 – Dégénérescence 2

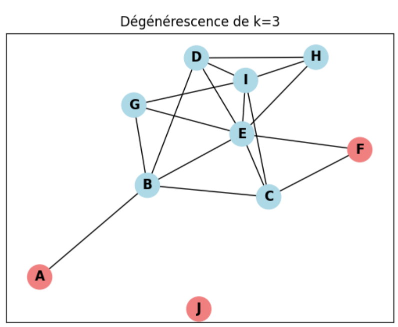

Figure 6 – Dégénérescence 3

La dégénérescence d'un graphe offre une mesure de sa structure interne, en mettant en évidence des sous-ensembles rassemblant des sommets. Cependant, malgré sa pertinence sur l'analyse des réseaux, la dégénérescence seule ne suffit pas toujours à capturer toutes les subtilités de la connectivité dans un graphe. Ainsi, dans le prochain chapitre, nous explorerons l'algorithme de Bron-Kerbosch, une méthode efficace pour identifier et explorer les cliques maximales dans un graphe, offrant ainsi une perspective plus approfondie sur ses motifs de connectivité.

Algorithme de Bron-Kerbosch

L'algorithme de Bron-Kerbosch est un algorithme récursif, il cherche dans un graphe G le nombre maximal de cliques. L'algorithme passe trois variables en paramètres, ceux-ci sont pour notre exemple, R, P et X. L'équation $P \cup X = \Gamma(R)$ indique que pour la capturer les sommets que l'on veut ajouter aux cliques maximales, il faut prendre l'union de P et X, ce qui donne tous les ensembles de tous les sommets qui ont été sélectionnés et ceux-ci sont tous connectés à tous les sommets de P. La valeur de P représente l'ensemble de sommets déjà sélectionnés dans l'algorithme, X représente un ensemble de sommets candidats à être ajoutés à l'ensemble sélectionné et $\Gamma(R)$ est l'ensemble des voisins de tous les sommets dans l'ensemble R.

Fonctionnement Bron-Kerbosh

L'explication du fonctionnement de l'algorithme, se déroule en plusieurs étapes. La condition d'arrêt met en condition si l'union des ensembles P et X sont vides, cela implique qu'il n'y a plus de sommets candidats à pouvoir être ajouter à la clique, ce qui implique que la clique actuelle est maximale, on retourne la valeur R comme clique maximale. La boucle principale, permet d'itérer sur chaque sommet v dans l'ensemble P, l'algorithme fonctionne de façon récursive avec les ensembles renouvelés. $P \cap \Gamma(v)$, ceci reprend tous les sommets de P qui sont voisins de v, $R \cup \{v\}$ c'est la clique qui est en cours de construction avec le somme v ajouté, $X \cap \Gamma(v)$, ce sont les sommets de X qui sont voisins de v. On retire par la suite v de l'ensemble complet P et on l'ajoute à l'ensemble X. L'algorithme se répète jusqu'à avoir parcouru tous les sommets de l'ensemble P.

Function BronKerbosch(P, R, X)

if $P \cup X = \varnothing$ **then**
    report R as a maximal clique
end if
for each vertex $v \in P$ **do**
    BronKerbosch($P \cap \Gamma(v), R \cup \{v\}, X \cap \Gamma(v)$)
    $P \leftarrow P \setminus \{v\}$
    $X \leftarrow X \cup \{v\}$
end for

Cette partie montre l'algorithme de Bron-Kerbosch à présent mettons ça en pratique sur un exemple bien concret. La section suivante montre un exemple dans lequel on parcourt l'algorithme pour les deux premières cliques afin de mieux comprendre son fonctionnement pas à pas.

Exemple Bron-Kerbosch

La figure 7 représente le graphique sur lequel va se dérouler notre exemple et sur lequel l'algorithme de Bron-Kerbosch va être exécuté.

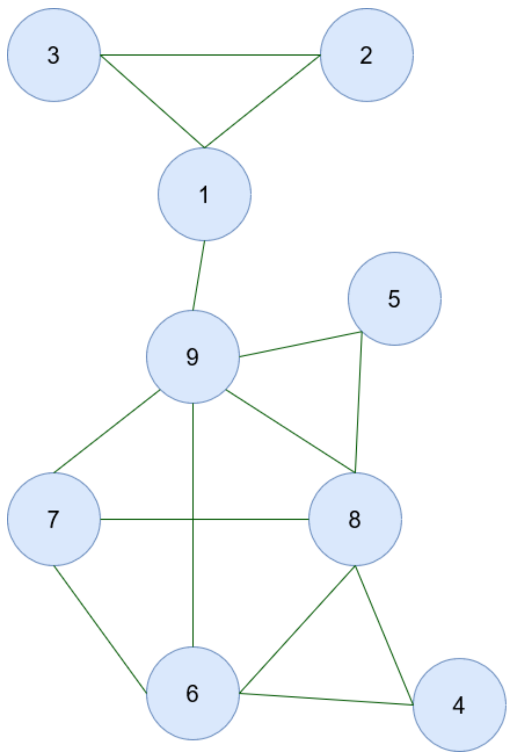

FIGURE 7 – Graphique Bron-Kerbosh

Dans un premier temps, v prendra la valeur de 1, c'est la première itération sur l'ensemble P. Celui-ci vaut $P = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, lors de la première récursivité de l'algorithme, les différentes valeurs en paramètres sont $P \cap \Gamma(v) = \{9, 2, 3\}$, $R \cup \{v\} = \{1\}$ et $X \cap \Gamma(v) = \{\}$. La valeur de v passe à 9, ce qui fait que $P \cap \Gamma(v)$ et $X \cap \Gamma(v)$ sont tous deux vides, ce qui nous donne notre première clique maximale sur la figure 8.

Ensuite, on va revenir à $v = 2$, ce qui implique que les valeurs de P et R changent également, $P \cap \Gamma(v) = \{3\}$, $R \cup \{v\} = \{1, 2\}$ et $X \cap \Gamma(v) = \{\}$. On passe à la valeur $v = 3$, ce qui a pour effet de rentrer dans la condition de sortie pour

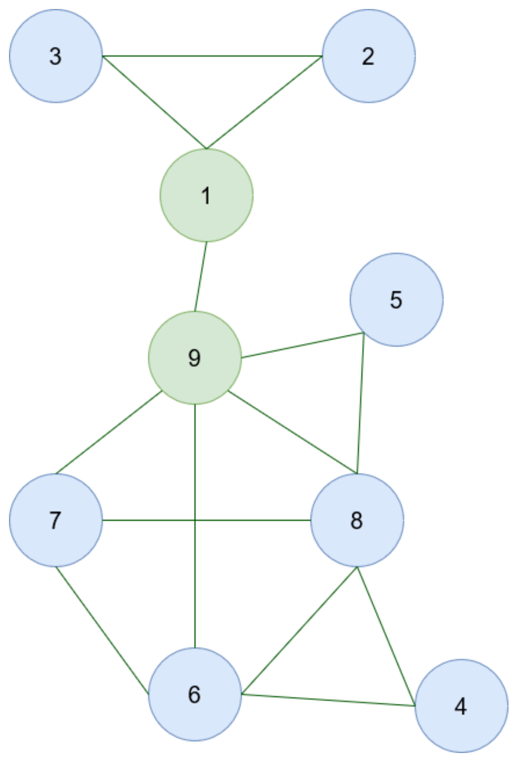

FIGURE 8 – Graphique Bron-Kerbosh 1-9

lesquelles P et X sont vides, ce qui donne notre deuxième clique maximale à la figure 9.

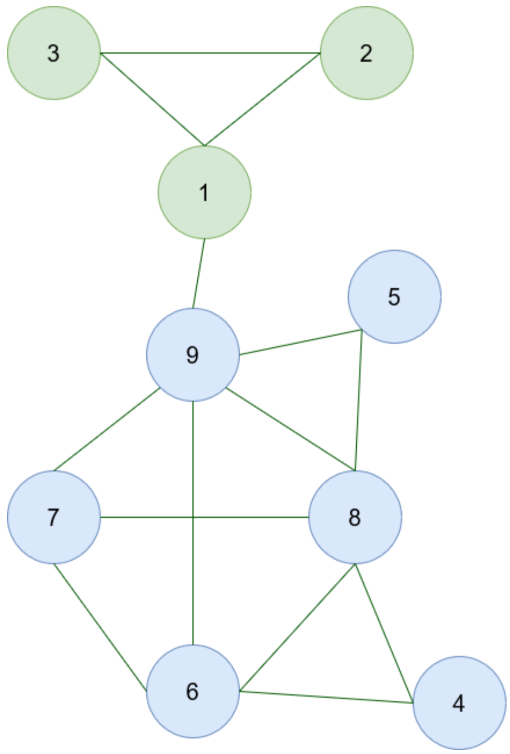

FIGURE 9 – Graphique Bron-Kerbosh 1-2-3

Au final, l'algorithme détecte six cliques maximales, ceux-ci sont les deux premiers pour lesquelles les étapes ont été démontrées, ainsi que les cliques : $R = \{8, 4, 6\}$, $R = \{8, 9, 5\}$, $R = \{8, 9, 6, 7\}$, $R = \{9, 6, 7\}$

Présentation de l'algorithme de l'article

Conclusion

Bibliographie

[1] Austin BUCHANAN et al. “Solving maximum clique in sparse graphs : an $O(nm+n^2d/4)$ algorithm for d-degenerate graphs”. In : *Optimization Letters* 7.7 (2013). Original Paper, p. 1513-1523. DOI : 10.1007/s11590-013-0698-2.

[2] Edmund R. PEAY. “Hierarchical Clique Structures”. In : *Sociometry* 37.1 (1974), p. 54-65.

[3] WIKIPÉDIA. “Degré (théorie des graphes)”. In : (20 septembre 2023). Consulté le 03 Février 2024. URL : https://fr.wikipedia.org/wiki/Degr%C3%A9_(th%C3%A9orie_des_graphes).