

A. Castillon : Compression de graphes en utilisant les quasi-cliques

Julien Baste, CRISAL, Lille, julien.baste@univ-lille.fr

Antoine Castillon, ENS de Lyon, Lyon, antoine.castillon@ens-lyon.fr

Mohammed Haddad, LIRIS, Lyon, mohammed.haddad@univ-lyon1.fr

Hamida Seba, LIRIS, Lyon, hamida.seba@univ-lyon1.fr

En 2020, Wang et al. [1], ont introduit une nouvelle méthode de compression sans perte d'information en utilisant les quasi-cliques. La compression d'un graphe se fait pour chaque quasi-clique de la façon suivante : on garde en mémoire tous les sommets qui la composent ainsi que les paires de sommets non adjacents. Pour des quasi-cliques de densité supérieure à 0.5 il est clair que cette représentation est moins coûteuse que de stocker toutes les arêtes.

Nous avons étudié cette méthode de compression en nous concentrant surtout sur la phase de détection des quasi-cliques. Nous nous sommes d'abord intéressés au Quasi-Clique Enumeration Problem (QCE), qui est le problème d'énumération de toutes les quasi-cliques d'un graphe. En nous basant sur l'algorithme Quick [2], nous avons proposé une nouvelle technique de précalcul permettant d'accélérer significativement la résolution de ce problème. L'énumération de toutes les quasi-cliques d'un graphe pouvant conduire à des redondances lors de la compression, nous avons ensuite étudié comment mieux adapter la phase de détection des quasi-cliques à la compression du graphe. Nous nous sommes donc intéressés à d'autres problèmes de détection de quasi-cliques où l'on impose une certaine diversité des résultats [3] ainsi qu'à leur application à la compression.

Références

- [1] Wang et al, *Dense Subgraphs Summarization : An Efficient Way to Summarize Large Scale Graphs by Super Nodes*, Intelligent Computing Methodologies (2020), 520-530.
- [2] Liu and Wong, *Effective Pruning Techniques for Mining Quasi-Cliques*, (2008), DOI : 10.1007/978 - 3 - 540 - 87481 - 2_3
- [3] Wang et al, *Redundancy-Aware Maximal Cliques*, Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2013), 122-130.