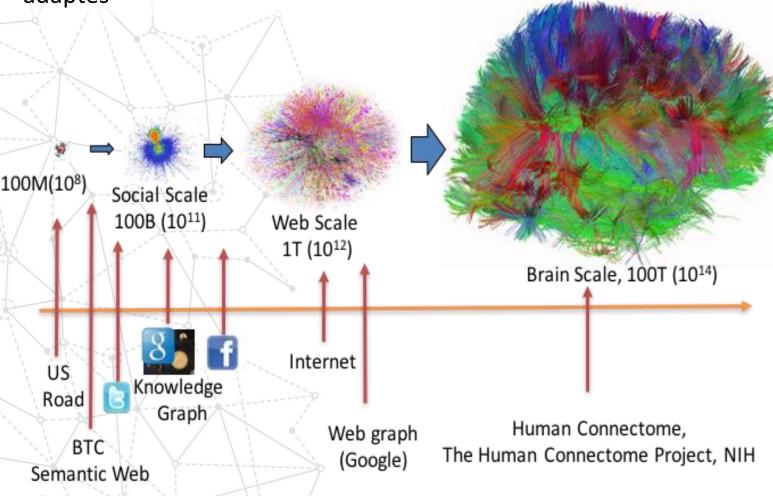
République algérienne démocratique et populaire الجمهورية الجزائرية الديمقراطية الشعبية Ministère de l'Enseignement supérieur et de la Recherche scientifique وزارة التعليم العالى والبحث العلمى



Étude et classification des méthodes de compression de graphes par extraction de motifs et les arbres k2-trees

Contexte et Problématique

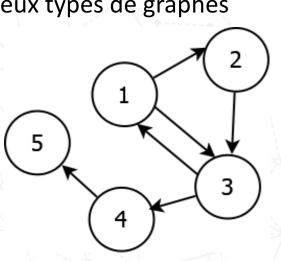
- La puissance des processeurs augmente plus vite que les capacités de stockage ce qui engendre un déséquilibre entre le volume des données qu'il est possible de traiter et de stocker.
- Capacité mémoire : les graphes ne peuvent plus résider en mémoire vue la quantité d'informations qui modélisent de nos jours.
- Les algorithmes classiques de traitement des graphes ne sont plus adaptés



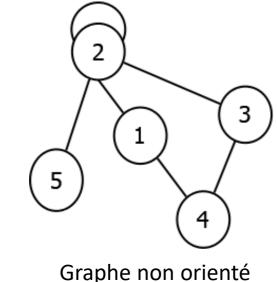
Comment gérer cette grande masse de données tout en permettant une exécution efficace des algorithmes classiques de traitement des graphes?

Aperçu général de l'état de l'art

Graphe : Un graphe G = (V, E) est un couple formé d'un ensemble de nœuds V = {v1, v2, ..., vn} et un ensemble de liens E = {e1, e2, ..., en}. Chaque lien e de l'ensemble E est défini par une paire de nœuds de l'ensemble V. c'est-à-dire e = (vi , vj). Nous distinguons deux types de graphes



Graphe orienté



Décompression

La compression des graphes : La compression de graphes est

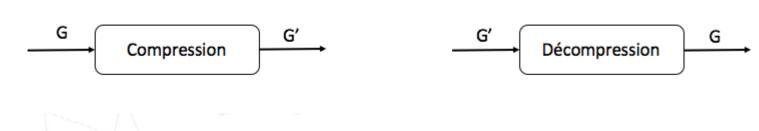
définie comme l'ensemble des méthodes et techniques permettant de réduire l'espace mémoire occupé par ces derniers tout en gardant la même signification que le graphe d'origine et ses composants les plus importants.

Deux types de compression existent :

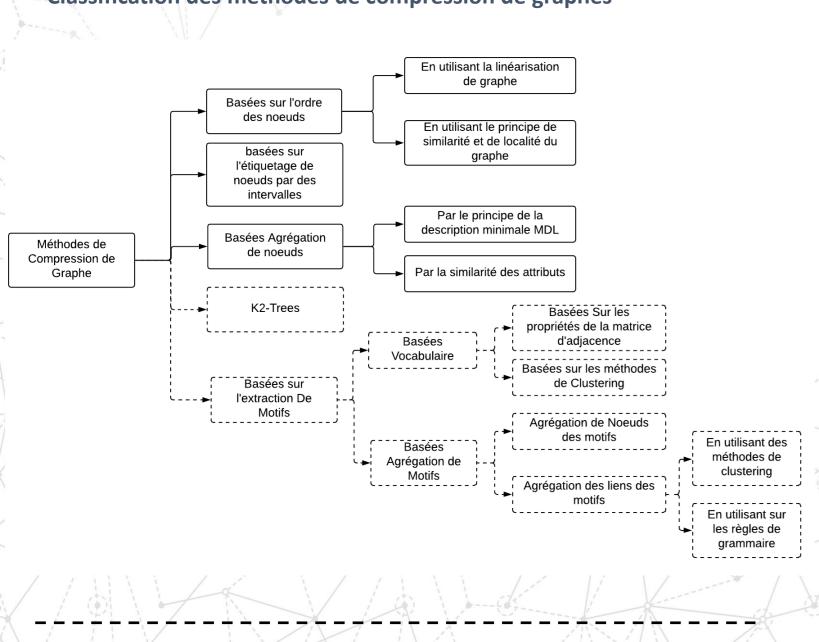
Compression sans perte :

Comparaison avec perte:

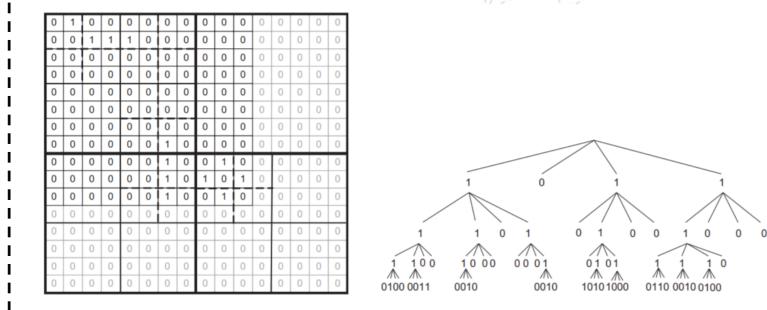
Compression



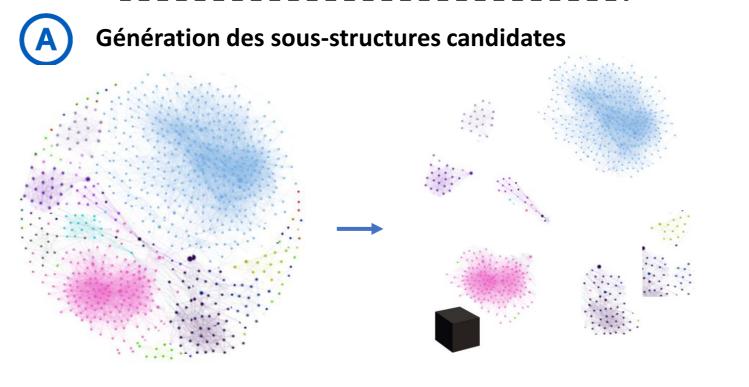
Classification des méthodes de compression de graphes

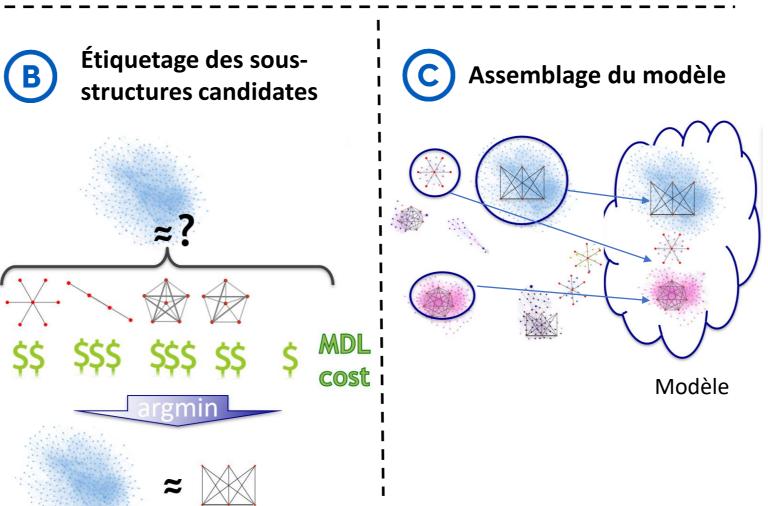


Les principaux travaux de Recherche:



K2-tree est une méthode de compression sans perte qui est appliquée sur des graphes statiques non pondérés. Elle est destinée pour les graphes du web. Elle exploite les zones vides dans la matrice d'adjacence.





VoG est une méthode de compression sans perte basée sur les méthodes de clustering, qui utilise le principe MDL pour choisir un ensemble de structures appelé modèle permettant d'obtenir une représentation succincte.

Comparaison entre les différentes Classes

Conclusion

La compression est une technique qui permet de manipuler les graphes de manière plus efficace.

- Les méthodes de compression basées sur l'extraction de motifs permettent de synthétiser le graphe à travers ses motifs les plus importants offrant une meilleur analyse et une visualisation plus claire.
- 2. Les méthodes de compression basées sur les arbres k2trees exploitent les propriétés de localité et de similarité de la matrice d'adjacence et permettent l'extraction de voisinage.

Perspectives

A court terme

- 1. Conception d'un outil de compression de graphe par extraction de motifs et k2-trees.
- 2. Implémentation de plusieurs méthodes existantes.
- 3. Evaluation des performances de chaque approche en utilisant des benchmarks connus.
- 4. Comparaison entre les méthodes de différentes classes.

A long terme

 Développement de nouvelles méthodes de compression compétitives en temps d'exécution et de taux de compression tout en se basant sur l'extraction de motifs et les arbres k2trees.

Références

- 1. Brisaboa, N. R., Ladra, S., & Navarro, G. (2009, August). k 2-trees for compact web graph representation. In *International Symposium on String Processing and Information Retrieval* (pp. 18-30). Springer, Berlin, Heidelberg.
- 2. Koutra, D., Kang, U., Vreeken, J., and Faloutsos, C. (2015). Summarizing and understan-ding large graphs. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 8(3):183–202.
- 3. Liu, Y., Safavi, T., Dighe, A., and Koutra, D. (2018a). Graph summarization methods and applications: A survey. *ACM Computing Surveys (CSUR)*, 51(3):62. 4.Álvarez-García, S., Freire, B., Ladra, S., and Pedreira, Ó. (2018). Compact and efficient representation of general graph databases. Knowledge and Information Systems, pages 1–32.
- 5.Shah, R. J. (2018). Graph compression using pattern matching techniques. arXiv preprint arXiv:1806.01504.
- 6.Liu, Y., Safavi, T., Shah, N., and Koutra, D. (2018b). Reducing large graphs to small super- graphs: a unified approach. Social Network Analysis and Mining, 8(1):17.

				Graphe En entrée				Type de Compression		Structure en sortie	
				Statique	Dynamique	Attribué	Étiqueté	Sans Parte	Avec Perte	Succinte	Structurelle
Méthodes basées sur les arbres k2-trees Méthode de Base (de Bernardo Roca, 2014) Att <i>k</i> 2-trees (Álvarez-García et al., 2018)			Х	Х	Х	Х	Х		Х		
Méthodes basées sur l'extraction de motifs	Basées	Basées sur les propriétés de la matrice d'adjacence GCUPM(Shah, 2018)		Х				Х		Х	
	Vocabulaire	Basées sur les méthodes de clustering VoG (Koutra et al. 2015) ConDenSE (Liu et al. 2018)		X	X			X		X	
		Agrégation des nœuds des motifs GraphZip (Rossi et Zhou, 2018)		Х			Х				Х
	1	Agrégation des liens des motifs	En utilisant des règles de grammaire gRepair(Maneth et Paterneck, 2018)	Х				X			X
			En utilisant les méthodes de clustreing DSM(Hernandez et Nvarro, 2014)	X				X			X

