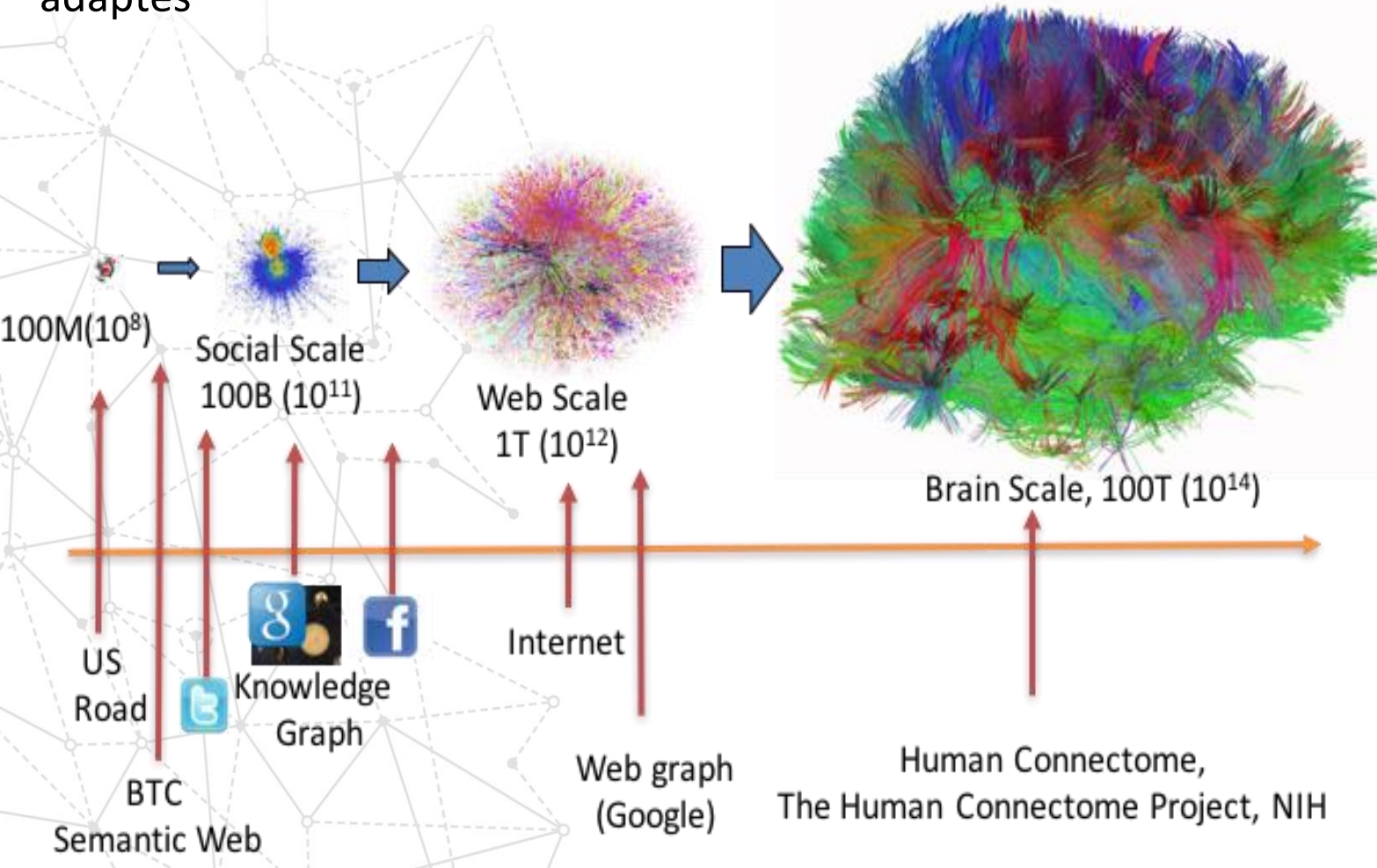


Étude et classification des méthodes de compression de graphes par extraction de motifs et les arbres k2-trees

Contexte et Problématique

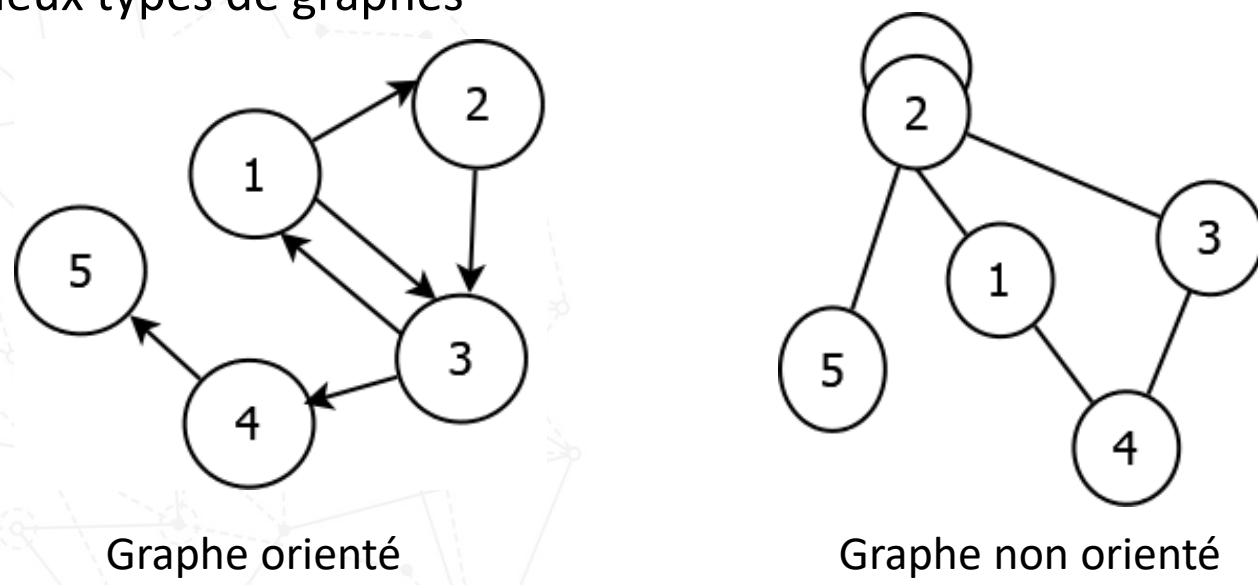
- La puissance des processeurs augmente plus vite que les capacités de stockage ce qui engendre un déséquilibre entre le volume des données qu'il est possible de traiter et de stocker.
- Capacité mémoire : les graphes ne peuvent plus résider en mémoire vue la quantité d'informations qui modélisent de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

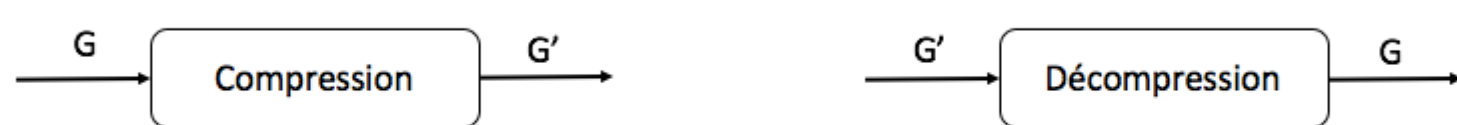
Graphe : Un graphe $G = (V, E)$ est un couple formé d'un ensemble de nœuds $V = \{v_1, v_2, \dots, v_n\}$ et un ensemble de liens $E = \{e_1, e_2, \dots, e_n\}$. Chaque lien e de l'ensemble E est défini par une paire de nœuds de l'ensemble V . c'est-à-dire $e = (v_i, v_j)$. Nous distinguons deux types de graphes



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 :

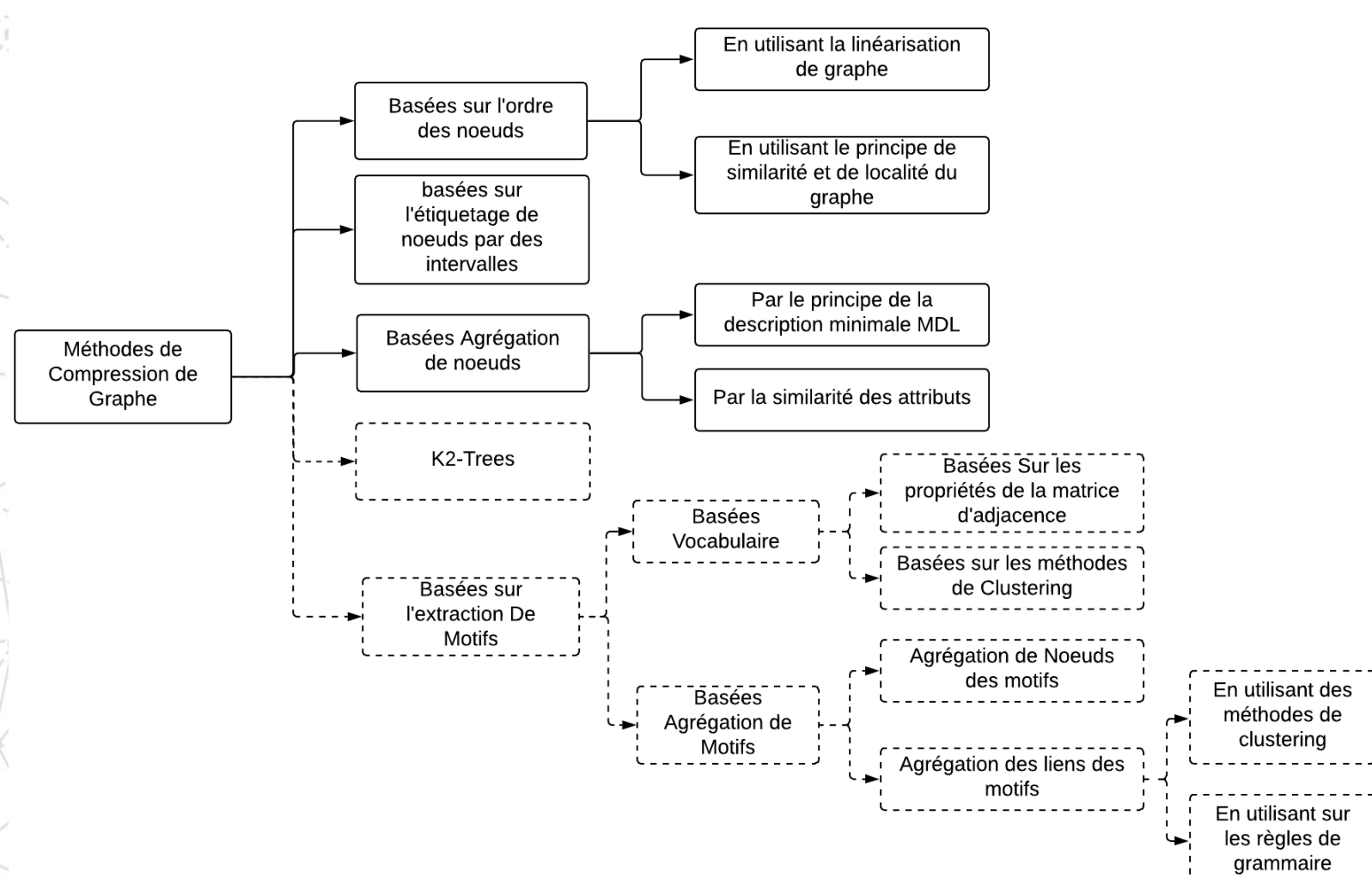
- Compression sans perte :



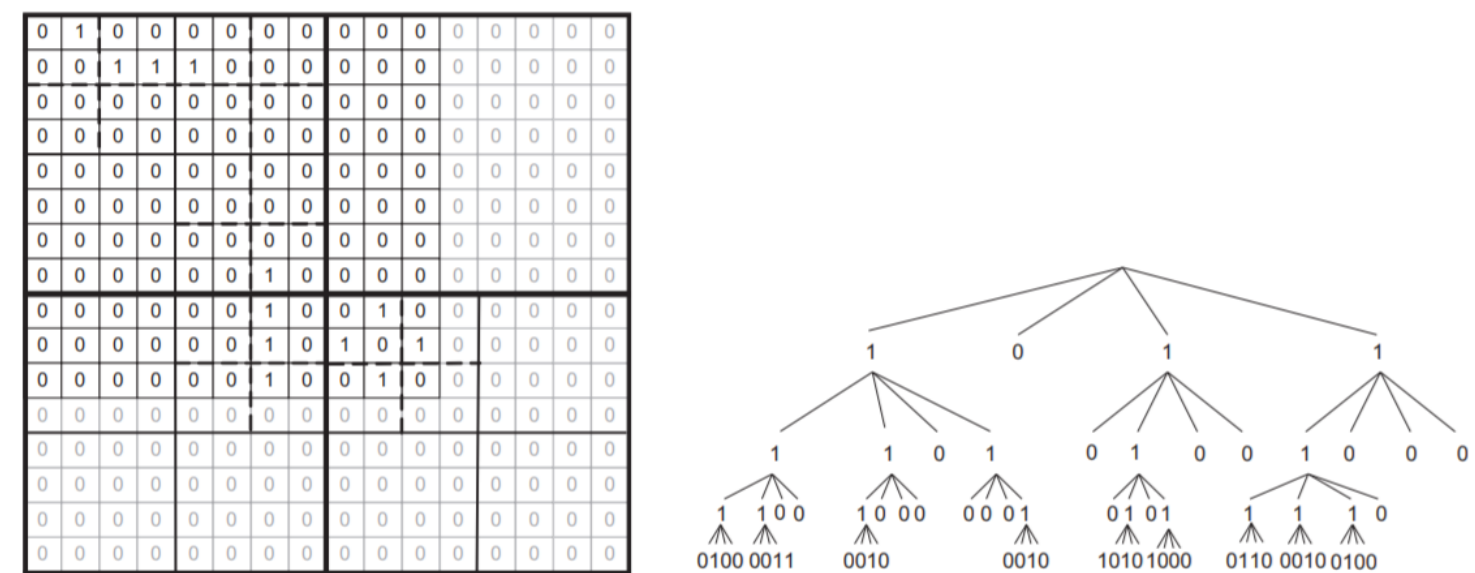
- Comparaison avec perte :



Classification des méthodes de compression de graphes

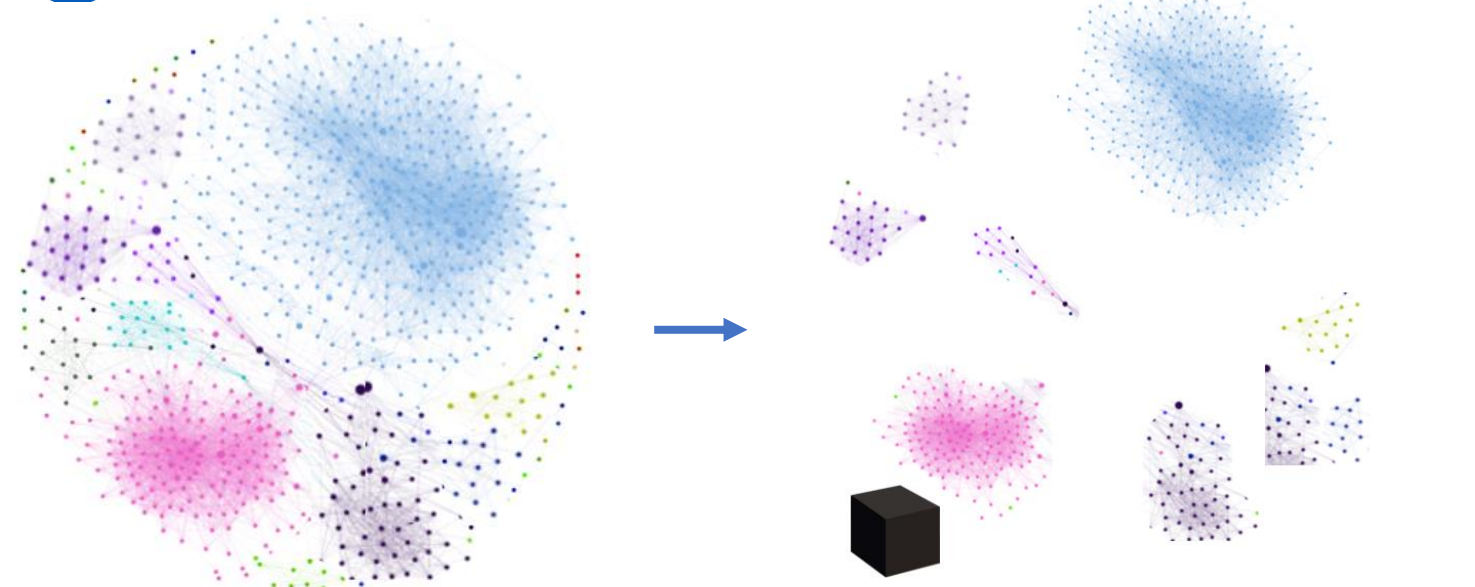


Les principaux travaux de Recherche :

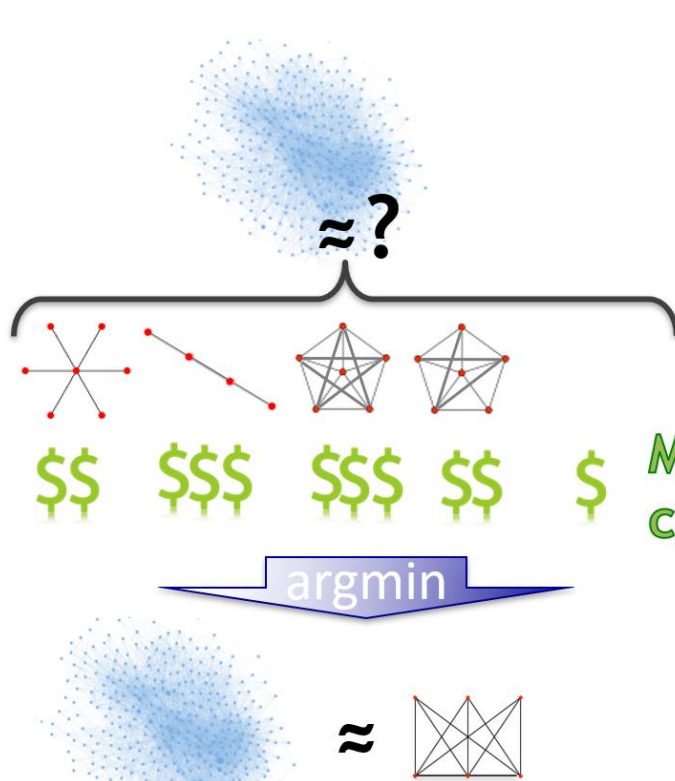


K2-tree est une méthode de compression sans perte qui est appliquée sur des graphes statiques non pondérés. Elle est destinée pour les graphes du web. Elle exploite les zones vides dans la matrice d'adjacence.

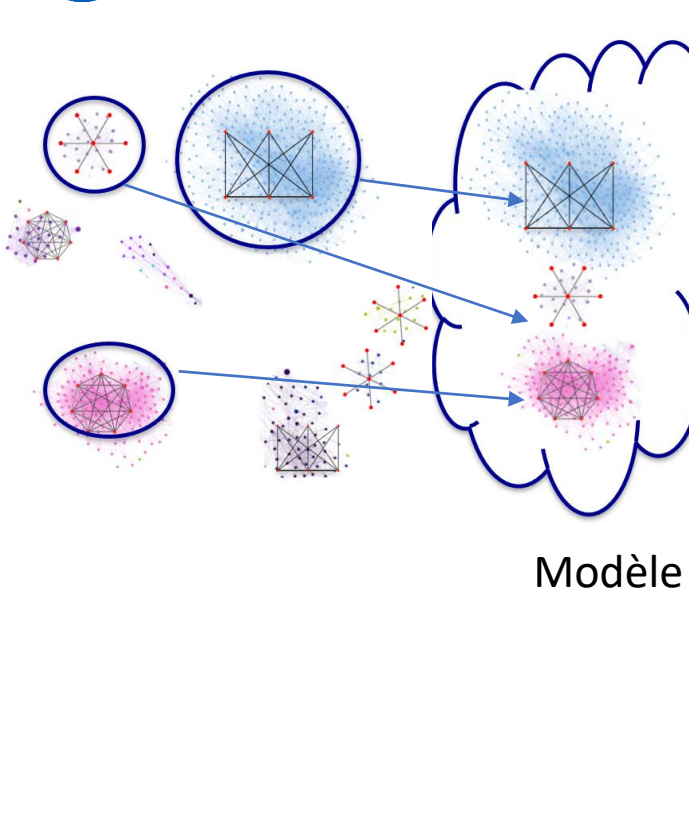
A Génération des sous-structures candidates



B Étiquetage des sous-structures candidates



C Assemblage du modèle



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

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

