

# Etude de graphes du Web

**Introduction :** Le but du projet est d'étudier différents paramètres de graphes issus du web. Typiquement, ces graphes sont grands en termes de nombre d'arêtes et de sommets, et leur étude est souvent complexe, car elle requiert un nombre important de calculs. Dans ce projet, on s'intéressera à différents graphes du Web. D'une part on étudiera des graphes générés aléatoirement modélisant des graphes réels du web. D'autres part on utilisera des graphes issus de la base de données des [grands graphes de Stanford](#), issus de différents réseaux sociaux. Typiquement, ces graphes sont représentés sous forme de fichier texte, dans lequel sont donnés leurs arêtes et sommets.

## Partie 1. Génération de grands graphes aléatoires.

**Objectif:** On cherchera à créer des graphes les plus grands possibles en terme de nombres de sommets. Il faudra ensuite les stocker dans un fichier texte, que vous utiliserez dans la deuxième partie du projet. On vous laisse le choix quant à la façon dont les graphes seront stockés. Par exemple, vous pourrez stocker leur liste d'adjacence ou bien simplement la liste des arêtes.

**Instructions:** On générera les graphes aléatoires de deux façons différentes.

Premièrement, étant donné un nombre fixe de sommet, chaque arête apparaît avec probabilité  $p=1/2$ . Ce modèle a été proposé par Edgar Gilbert.

La seconde famille de graphes aléatoires que l'on vous demande de générer sont les graphes de Barabási-Albert. Ces graphes sont construits itérativement de la façon suivante. On considère un paramètre  $m$  (ici on prendra  $m=2$ ). Initialement le graphe est un graphe triangle (une clique à trois nœuds). On ajoute des nœuds un par un. Supposons que les nœuds du graphe existant sont numérotés. Un nouveau nœud sera connecté au nœud  $i$  avec probabilité  $\text{degré}(i) / \sum \text{degrés}$ . On créera au plus  $m$  nouvelles telles arêtes. Par exemple, lorsque l'on ajoutera le quatrième nœud au triangle initial, il sera connecté avec probabilité  $2/3$  à chacun des nœuds du graphe existant.

## Partie 2. Etude des graphes.

On vous demande d'étudier certains paramètres pour les graphes générés dans la première partie du projet, ainsi que pour certains graphes issus de la base de données des [grands graphes de Stanford](#), qui se trouvent dans le dossier intitulé "Grand graphes". Veuillez noter que tous ces graphes ne sont

pas stockés dans le fichier texte de la même façon. Ainsi vous aurez probablement besoin de modifier le fichier texte pour exécuter vos algorithmes.

**Instructions:** Pour les graphes de la première partie du projet, vous calculerez les paramètres sur différentes tailles de graphes (une dizaine au maximum), l'objectif étant de voir la taille maximum du graphe (en terme de sommets) sur lequel il est possible de calculer les paramètres. Vous calculerez également la valeur des paramètres pour les graphes issus de la base de donnée Stanford. On attend que vos résultats soient stockés dans un tableau de la façon suivante:

Type de graphe	paramètre 1	paramètre 2	etc.
Graphe Edgar Gilbert			

Voici la liste des paramètres à calculer pour chacun des graphes.

1. Nombre de sommets.
2. Nombre d'arêtes.
3. Degré maximal.
4. Degré moyen (somme des degrés divisée par le nombre de sommets).
5. La distribution des degrés, sous forme de graphique. En abscisse, vous mettrez le degré et en ordonnée sa fréquence d'apparition dans le graphe.
6. (**bonus**) Le diamètre du graphe (le plus long plus court chemin entre n'importe quelle paire de sommets). Vous décrirez l'algorithme utilisé, ainsi que sa complexité.

**Format du rendu:** On attend un rapport pdf contenant les résultats attendus dans la partie 2. Votre code devra être écrit en C++ ou python. Il devra être propre et commenté. Un readme expliquant la compilation/exécution est obligatoire. Vous ferez très attention au format des graphes issus de la base de donnée Stanford, et les modifierez au besoin.