La question que nous essayons de répondre est :

Est-il possible de générer un graphe de réseau social aléatoire (et volumineux) qui est statistiquement indiscernable du graphe d’un réseau social réel comme Facebook par exemple ?

La réponse courte est oui, mais il existe plusieurs façons différentes de le faire. Nous allons vous présenter deux approches sur les techniques de génération, la première sera basée sur des graphes aléatoires et la seconde sera un processus génératif récursif comme une fractale.

## Quelques définitions :

Avant de détailler les différentes approches, parlons d’abord des graphes complètements aléatoires. La particularité de ces graphes est que chaque nœud du graphe a une probabilité d’être lié à tout autre nœud, cette probabilité détermine la densité du graphe telle que si le graphe généré sera un graphe complet, et si le graphe généré sera un graphe isolé (sans liens). Ce modèle de génération est connu sous le nom de **Erdős–Rényi.**

Les figures suivantes montrent la différence entre des graphes de 20 nœuds générés en fonction de la valeur de :

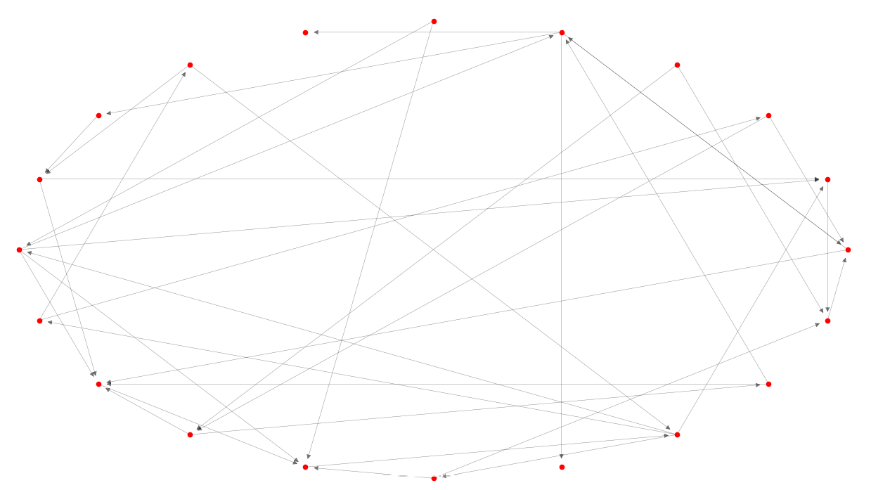


Figure : Graphe aléatoire p=0,1

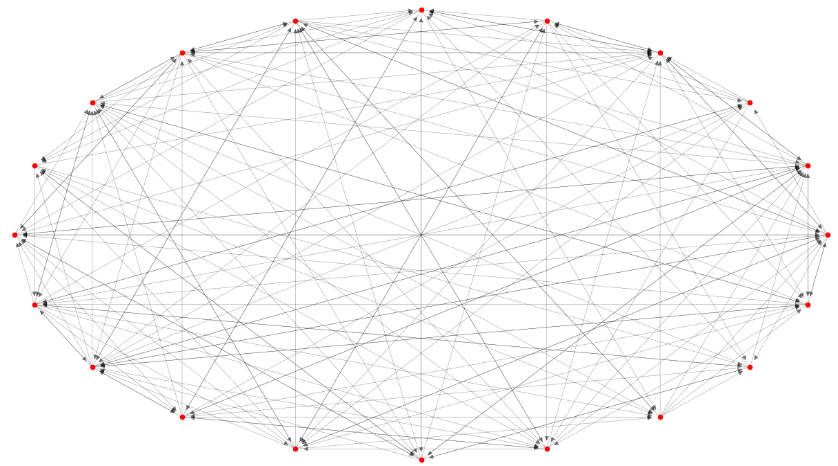


Figure : graphe aléatoire avec p=0.4

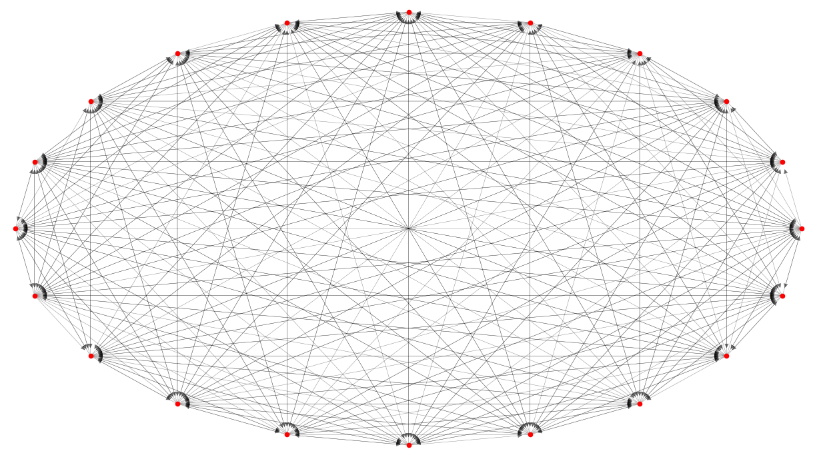


Figure : graphe aléatoire avec p=0.8

Nous palerons souvent de distribution de degrés, c’est une courbe de nombre de nœuds en fonction de leurs degrés, la courbe nous montre la concentration des degrés des nœuds dans le graphe.

La distribution des degrés du graphe **Erdős–Rényi** (graphe ER) suit la loi normale, ça veut dire que la majorité des nœuds possèdent un degré proche du degré moyen du graphe.

Voici un exemple d’un graphe ER de 1000 nœuds et :

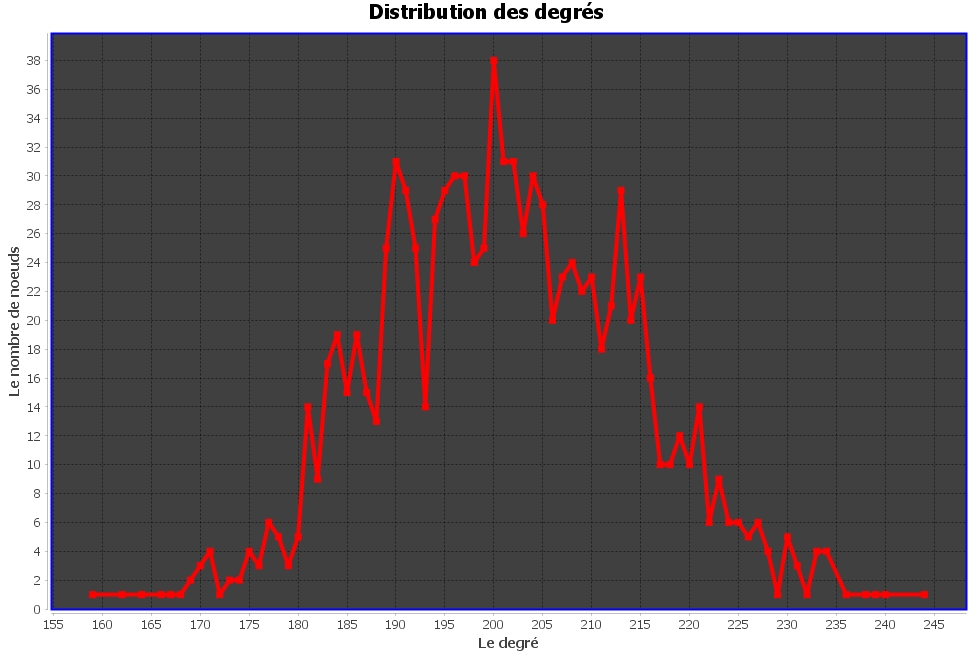


Figure : distribution des degrés d'un graphe ER de 1000 nœuds et p=0.1

Le graphe ER précèdent est loin de ressembler à un graphe d’un réseau social, car selon Albert-László Barabási et Réka Albert qui sont des experts dans le domaine de l’analyse des réseaux, la distribution des degrés dans un réseau social suit une loi de puissance où , *" Une propriété commune a beaucoup de grands réseaux est que les connectivités de vertex suivent une distribution de loi de puissance sans échelle "* (1), ça veut dire que dans un réseau réel (en particulier des réseaux sociaux en ligne) le degré de distribution est très asymétrique contrairement au graphe ER, il y aura beaucoup de nœuds ayant un degré bas alors qu'une petite fraction (mais significative) de nœuds auront un degrés plus élevé.

La courbe de la distribution des degrés d'un réseau social réel ressemblera à la courbe suivante :

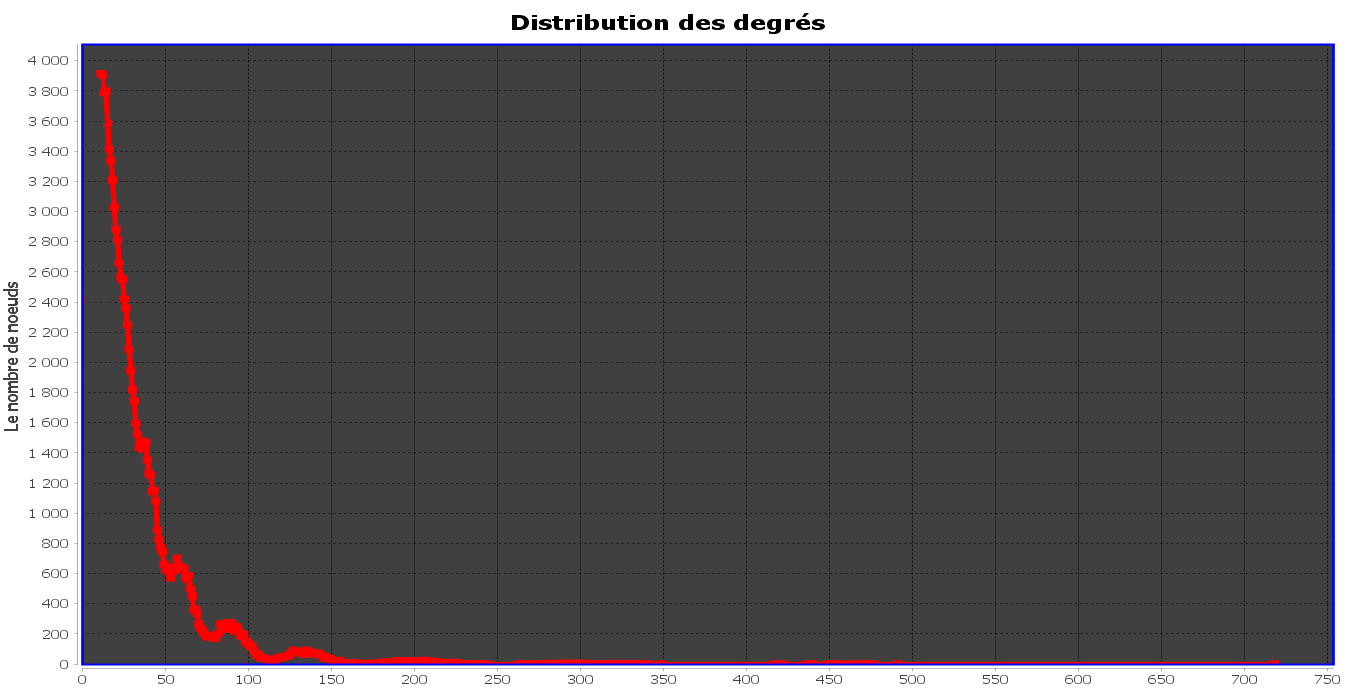


Figure : la distribution des degrés d'un graphe de 100000 nœuds et degré moyen 20

## Première approche : réseau social à partir d'un graphe aléatoire

# L'idée général consiste à construire un graphe d'une façons itératif (ajouter un nœud à la fois), en partant d'un graphe ER tout en essayant de se rapprocher le plus de la structure d'un graphe de réseau social à l'aide de la distribution des degrés. Ce modèle de génération est connu sous le nom de modèle de Barabàsi-Albert portant les noms des deux chercheurs Albert-László Barabási et Réka Albert.

# Fonctionnement :

# A partir du graphe ER initiale, on rajoute un nouveau nœud , ce dernier aura une probabilité d'avoir un lien avec tous les autres nœuds suivant la formule suivante :

# Où est degrés du nœud , et est la somme des degrés du graphe initial, et un paramètre pour contrôler la densité du graphe.

# Pour l'itération suivante le graphe initial sera le graphe résultant de l'itération précédente. On répète cette opération autant de fois qu'il faut pour arriver au nombre de nœuds souhaité.