0.1.1 Construction de graphe

[Modèle 1] On se donne $V = 1, \cdots, 2N$ un ensemble, $p, p_A, p_B \in [0,1]$ des probabilités et on suppose qu'on dispose de $A \subset V$, tel que $A = N$ construit aléatoirement, on note alors $B = VA$. On alors: $V = A \sqcup B$ $B = A = N$.

On définit $\sigma : V \longrightarrow -1, 1$ la fonction d'étiquetage des sommets telle que:

$$\forall v \in V, \sigma(v) = \begin{cases} 1 & siv \in A \\ -1 & siv \in B \end{cases}$$

On construit alors les arêtes de E de la façon suivante pour tout $(u, v) \in V^2$.

$$[(u,v) \in E] = \begin{cases} p & si\sigma(u) \neq \sigma(v) \\ p_A & siu, v \in A \\ p_B & siu, v \in B \end{cases}$$

Dans la suite on notera $G(N, p_A, p_B, p)$, G est un graphe aléatoire construit selon ce modèle.

0.1.2 Exemples

0.3 [scale=0.5]ig-plots/large$_inter_proba.pdf$

Figure 1: $p_A = p_B = 0.8, p = 0.1$

0.3 [scale=0.5]ig-plots/large$_small_smaller.pdf$

Figure 2: $p_A = 0.9, p_B = 0.3, p = 0.05$

0.3 [scale=0.5]ig-plots/large$_largenoise.pdf$

Figure 3: $p_A = p_B = 0.9, p = 0.5$

Figure 4: Exemples de graphes générés par le modèle avec différents paramètres

0.2 Problème de reconstruction exacte

On suppose maintenant que l'on dispose d'un graphe $G(N, p_A, p_B, p)$ et que A et B étaient les deux communautés originales et σ la fonction d'étiquetage associée.

[Problème de reconstruction exacte] On cherche à reconstruire σ en l'estimant à partir uniquement de l'observation du graphe. Le problème revient donc à construire $\hat{\sigma}$ tel que:

$$[\hat{\sigma} = \pm\sigma] N \to \infty \longrightarrow 1$$

Cela revient à retrouver A et B à renommage près.

Nous allons examiner une solution particulière de ce problème.

0.3 Bissection minimale

[Bissection et coupe] Soit $G = (V, E)$ un graphe une bissection de G est une partition des sommets en deux ensembles A, B.

On parle alors de la coupe $A : B$ et de la bissection $A : B$ si A et B sont de tailles égales. On note de plus $A : B$ l'ensemble des arêtes dont une extrémité est dans A et l'autre dans B.

Par la suite on utilisera la notation $A : B$ pour introduire la coupe A On parlera par la suite de la bissection ou de la coupe $A : B$ suivant si A et B sont de tailles égales

[Flot] Soient $G = (V, E)$ et $A, B \subset V$ disjoints. On appelle flot entre A et B le nombre d'arêtes (ou la somme des capacités des arêtes)

On note cette valeur $A : B$. En effet, dans le cas d'un graphe non pondéré, le flot correspond au cardinal de $A : B$.

[Ordre sur les bissections] Soit $G = (V, E)$, et $A : B$, $A' : B'$ deux bissections de G. On dit que $A : B \preceq A' :' B$ si et seulement $A : B \leq A' : B'$.

C'est à dire une bissection est plus petite qu'une autre si le flot, ie la capacité des arêtes allant d'un ensemble à l'autre de la bipartition, est plus petit que le flot de l'autre

Dans ces conditions la − ou les bissections − minimales sont les bissections minimales pour $\prec$, *ie.* toutes les autres bissections envisageables ont un flot au moins supérieur au leurs.

Il se trouve que le cadre du modèle 0.1.1 précédemment énoncé, la bissection minimale est presque sûrement solution du problème de reconstruction exacte. Dans la section suivante nous allons voir la preuve de ce résultat avant d'analyser la complexité du problème de a bissection minimale d'un graphe.

Nous allons étudier le problème de reconstruction exacte de communautés dans un graphe. Tout d'abord nous définirons un modèle de graphe aléatoire modélisant les graphes induits par des communautés pré-établies puis nous donnerons une méthode reconstruction exacte avec grande probabilité de ces communautés avant d'en étudier en détail la classe de complexité algorithmique. Plus précisément, nous montrerons qu'une bissection minimale du graphe est solution du problème de reconstruction exacte avec grande probabilité mais que la recherche d'une telle bissection est un problème $NP-Complet$ **0.1 Modèle**