Structures et algorithmes aléatoires
TD7

14 novembre 2014

Exercice 1 Routage dans les graphes petit-monde

En 1967, le sociologue Stanley Milgram publia les résultats d'une de ses expériences. Il demanda à des individus de transmettre chacun une enveloppe à un autre individu, avec comme information sur le destinataire, son nom, son adresse et sa profession. L'enveloppe ne pouvait pas être envoyée directement, mais passer par des connaissances. Il apparut qu'une grande quantité des enveloppes arrivèrent à destination et que le nombre de d'intermédiaires était d'au plus 6 dans la plupart des cas. C'est ce qu'on appelle le phénomène *petit monde*. Dans les graphes d'Erdős-Rényi, cela se traduit par le fait que le diamètre d'un graphe soit logarithmique en le nombre de sommets.

On s'intéresse ici au routage dans des graphes qui ont la propriété d'avoir un petit diamètre (ce que l'on admet dans cet exercice), et plus particulièrement au modèle de Kleinberg : les sommets du graphes sont sur une grille $\{1,\dots,m\}\times\{1,\dots,m\}$. Deux sommets sont voisins si $|u,v| = 1$ (où $|\cdot|$ est la distance L_1). On ajoute à cette grille des *raccourcis*, q par sommets. Un sommet u a un raccourci vers v avec probabilité $|u-v|^{-\alpha}/\sum_{w\neq u}|u-w|^{-\alpha}$.

On s'intéresse au cas où $\alpha = 2$. Et on considère le routage glouton suivant : à chaque pas, on va vers le voisin le plus proche (au sens L_1) de la destination finale. On fixe u une source, v une destination, et $u(t)$ est le sommet atteint après t pas. On dit que $u(t)$ est en phase j à la date t si $2^j < |u(t)-v| \le 2^{j+1}$. On note $T_{\text{alg}}(u,v)$ le nombre d'étapes pour aller de u à v avec cet algorithme.

1. Quelle est la probabilité que $u(t+1)$ appartienne une meilleure phase que $u(t)$? Montrer que cette probablité est supérieure à $1/72(1+\log(2m))$.
2. En déduire que $\mathbb{E}\big(T_{\text{alg}}(u,v)\big) = O\big(\log(n)^2\big)$.

On s'intéresse maintenant que cas où $\alpha \neq 2$ et nous allons montrer que cet algorithme glouton n'est plus efficace.

3. Montrer, dans le cas où $\alpha < 2$ que $\mathbb{E}\big(T_{\text{alg}}(u,v)\big) = \Omega\big(m^{\frac{2-\alpha}{3}}\big)$. On pourra pour ce faire s'intéresser au dernier raccourci pris lors d'un routage de longueur t et à la distance de ce dernier raccourci à v.
4. Montrer, dans le cas où $\alpha > 2$ que l'algorithme de routage de u vers v s'effectue en moyenne en $\mathbb{E}\big(T_{\text{alg}}(u,v)\big) = \Omega\big(|u-v|^{\gamma}\big)$ avec $\gamma = (\alpha-2)/(\alpha-1)$. On pourra tout d'abord calculer le probabilité d'avoir un raccourci de longueur au moins d à partir d'un sommet, puis borner la probabilité qu'on ne puisse pas avoir de routage en t étapes pour des sommets distants d'au moins $td+1$.

Exercice 2 Filtre de Bloom

Soit U un univers (ensemble fini de grand cardinal). On cherche à représenter un sous-ensemble S de U à m éléments. Pour ce faire, on utilise k fonctions de hachage parfaites dans un ensemble à n élément : on a des fonction $h_i : U \to [0,\dots,n-1]$ telles que pour tous $u \in U$ et $\ell \in [0,n-1]$, $\mathbb{P}(h_i(u) = \ell) = 1/n$, et les valeurs $h_i(u)$ sont mutuellements indépendantes.

On considère maintenant A un tableau à n cases, $A[0],\dots,[n-1]$. Pour tous $i \in [1,k]$, $s \in S$, on pose $A[h_i(s)] = 0$ et $A[\ell] = 0$ sinon.

Si l'on fait une requête sur un élément u, on retourne vrai si pour tout $i \in [1,k]$, $A[h_i(u)] = 1$. Si l'élément est bien dans S, la réponse est toujours correcte. Par contre, on peut avoir de faux positifs (un élément n'est pas dans S, mais on retourne vrai tout de même).

1. Quelle est la probabilité p que $A[\ell] = 0$? Quelle est la probabilité d'avoir un faux positif ?

Quand n est grand, p est proche de $e^{-km/n}$. On fixe donc pour la suite $p = e^{-km/n}$.

2. Pour m et n fixés, donner alors le meilleur k possible (on pourra dans un premier temps s'affranchir de la contrainte « k est un entier »).

On reprend $p' = (1 - 1/n)^{km}$, la probabilité que $A[\ell] = 0$. Soit X le nombre de cases ℓ telles que $A[\ell] = 0$.

3. Montrer que $\mathbb{P}(X - np' > \epsilon n) \leq e\sqrt{n}e^{-n\epsilon^2/3p'}$.