TD n°09

4 avril 2018

Exercice 1 - Filtre de Bloom

On souhaite vérifier que des mots de passe ne sont pas trop faciles à cracker en maintenant un dictionnaire de mots inacceptables. Quand un utilisateur rentre un mot de passe, notre application doit tester son appartenance à l'ensemble des mots inacceptables.
Un filtre de Bloom consiste en un tableau de n bits, $A[0]$ to $A[n-1]$ initialement mis à 0. Le filtre de Bloom utilise k fonctions de hachage aléatoires indépendantes $h_1, \ldots, h_k$. On fait l'hypothèse que ces fonctions de hachage associe à chaque élément de l'univers un nombre aléatoire uniforme dans $\{0, \ldots, n-1\}$. Soit $F = \{f_1, \ldots, f_m\}$ l'ensemble de mots de passe inacceptables.
L'étape de pré-processing est la suivante : Pour chaque élément $f_j \in F$, les bits $A[h_i(f_j)]$ sont mis à 1 pour tout $1 \leq i \leq k$. Observez qu'une entrée peut être mise à 1 plusieurs fois ; mais seulement la première affectation a un effet.
Pour vérifier si un élément x est dans F, on teste que toutes les entrées $A[h_i(x)]$ sont à 1 (pour tout $1 \leq i \leq k$). Si ce n'est pas le cas, on conclut que $x \notin F$. Il est facile de voir qu'on ne peut pas avoir de faux négatifs. Mais il peut y avoir des faux positifs : On peut avoir $A[h_i(x)] = 1$ (pour tout $1 \leq i \leq k$) et avoir tort de conclure que $x \in F$.

1.1 Soit X le nombre de positions dans A qui contiennent un 0 après pré-processing. Que vaut $\mathbf{E}\{X/n\}$?

1.2 Pour simplifier, supposons que $X = pn$ pour $p = e^{-km/n}$. Quelle est la probabilité P d'un faux positif ? Choisissez k pour minimiser P. Quelle valeur de p minimise P ?

1.3 Reconsidérez (et justifiez) notre hypothèse $X = pn$. Utilisez l'approximation de Poisson pour borner $\mathbf{P}\{|X - np| \geq \varepsilon n\}$.

Exercice 2 - Nombre chromatique

Le nombre chromatique $\chi(G)$ est le plus petit entier naturel χ tel qu'il est possible de colorer les sommets de G avec χ couleurs sans avoir les deux extrémités d'une arête de la même couleur. On pourrait penser que si le graphe n'a pas de cycles courts, alors $\chi(G)$ est borné. Cela s'avère pourtant faux.
Prouvez que pour tout entier $k \geqslant 2$, il y a un graphe G sans triangle de nombre chromatique k.

Exercice 3 - Graphe connexe

Un graphe non orienté est dit *connexe* s'il existe un chemin entre toute paire de sommets. Montrez que si $p = (2+\varepsilon)\log n/n$ pour un $\varepsilon > 0$, alors la probabilité qu'un graphe tiré aléatoirement de $G_{n,p}$ est connexe tend vers 1 quand n tend vers l'infini.

Exercice 4 - Deuxième lemme de Borel-Cantelli

Soit $(A_i)_{i=1}^{\infty}$ une famille d'événements mutuellement indépendants, et $\Sigma_{i=1}^{\infty}\mathbf{P}\{A_i\} = \infty$.
Montrez que $\mathbf{P}\{\text{une infinité de } A_i \text{ se réalise}\} = 1$.

Exercice 5 - Nombre normaux

Un nombre est dit *normal* si son développement dans toute base b possède toute séquence de longeur ℓ avec même fréquence $b^{-\ell}$.
Montrez que presque tout nombre réel est normal.