streaming algorithm Written Assignment $\sharp 1$

0656124 劉承順

1. $Pr[some\ s_i\ in\ S\ has\ rank\ r_A(s_i) \in [t-\sqrt{n}, t+\sqrt{n}]]$
   $= 1 - Pr[\bigcap_{s_i \in S}\ r_A(s_i) \notin [t-\sqrt{n}, t+\sqrt{n}]]$
   $= 1 - \prod_{s_i \in S} Pr[r_A(s_i) \notin [t-\sqrt{n}, t+\sqrt{n}]]$ (because s_i's are chosen independently)
   $Pr[r_A(s_i) \notin [t-\sqrt{n}, t+\sqrt{n}]] < \frac{n-\sqrt{n}}{n}$ for all $s_i \in S$ because $|[t-\sqrt{n}, t+\sqrt{n}]| \geq \lfloor\sqrt{n}\rfloor + 1$ for all t.
   $\prod_{s_i \in S} Pr[r_A(s_i) \notin [t-\sqrt{n}, t+\sqrt{n}]] < (1 - \frac{1}{\sqrt{n}})^k$
   because $k = \Omega(\sqrt{n})$, $\exists c > 0$ and n_0 such that $n > n_0 \Rightarrow k > c\sqrt{n}$.
   consider function $f(x) = (1 - \frac{1}{x})^x$ is increasing function when $x > 1$, and $\lim_{x\to\infty}(1 - 1/x)^{cx} = e^{-c}$.
   so $n > n_0 \Rightarrow k > c\sqrt{n} \Rightarrow (1 - \frac{1}{\sqrt{n}})^k < (1 - \frac{1}{\sqrt{n}})^{c\sqrt{n}} < e^{-c}$.
   $n <= n_0 \Rightarrow (1 - \frac{1}{\sqrt{n}})^k < (1 - \frac{1}{\sqrt{n}}) < (1 - \frac{1}{\sqrt{n_0}})$.
   then pick $\delta = min\{1 - e^{-c}, \frac{1}{\sqrt{n_0}}\}$, the proof is finished.

2. S is a simple graph of average degree n/3.
   for all nodes n_i, $0 \leq deg(n_i) \leq n - 1$.
   let x =ratio of nodes of degree at least n/4, $x = \frac{|\{n_i | deg(n_i) \geq \frac{n}{4}\}|}{n}$, $1 - x$ =ratio of nodes of degree less than n/4.
   we have $\frac{n}{3} \leq (n-1)x + \frac{n}{4}(1-x)$
   $(\frac{3}{4}n - 1)x \geq \frac{n}{12}$
   $x \geq \frac{1}{9}$ (consider $n \geq 4$)
   $|\{n_i, deg(n_i) \geq \frac{n}{4}\}| \geq \frac{1}{9}n$
   so $|\{n_i, deg(n_i) \geq \frac{n}{4}\}| = \Omega(n)$.

3. Let $X = |S_n|$, $X_k = 1$ if $k \in S$(with probability 1/k), $X_k = 0$ if $k \notin S$.
   we have $X = \Sigma X_k$.
   Let$\mu = E[X]$,$\mu = \Sigma_{k\in[1,n]} \frac{1}{k} = \Theta(logn)$.
   By Chernoff Bounds, we have
   $Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ for all $0 < \delta < 1$

$Pr[|S_n| = \Theta(logn)] \geq Pr[|X - \mu| \leq 0.5\mu]$
$\geq 1 - 2e^{-\mu 0.25/3} \geq 1 - \frac{1}{n^{\Omega(1)}}$

4. Let G_{abc} be a vertex-induced subgraph of G with three different nodes $a, b, c \in V(G)$.
   Pr[G_{abc} is triangle-free]$= 1 - (\frac{c}{n})^3$
   Pr[G is triangle-free]=Pr[$\wedge_{G_{abc} \subseteq G}(G_{abc}$ is triangle-free)]
   $\geq (1 - (\frac{c}{n})^3)^{\binom{n}{3}}$ (because "G_{abc} is triangle-free" is monotone decreasing.)
   $\geq (1 - (\frac{c}{n})^3)^{n^3}$ if $n > c$
   because $(1 - (\frac{c^3}{x}))^x$ is increasing function when $c > 0$,
   pick $n_0 > c$ and $\delta = (1 - (\frac{c}{n_0})^3)^{n_0^3}$
   $n > n_0 \Rightarrow (1 - (\frac{c}{n})^3)^{n^3} > \delta \Rightarrow$ Pr[G is triangle-free]$\geq \delta$

5. (a) 一般解法 (不限制 space)：答案的兩個正方形在座標圖上有兩種排法–左上加右下, 或左下加右上。因為任一個正方形的上邊必須剛好切過最高點 (y 值最大的點) 或是下邊必須剛好切過最低點。左邊必須切過 x 最小點，或是右邊必須切過 x 最大點。兩種排法分別找最小的正方形邊長，較小的則為解答。找極值:
   max{x|(x,y)$\in P\} = x_{max}$ (P is the point set)
   min{x|(x,y)$\in P\} = x_{min}$
   max{y|(x,y)$\in P\} = y_{max}$
   min{y|(x,y)$\in P\} = y_{min}$
   假設目前考慮兩個正方形是左下加右上的排法，令 $d =$ 最小的正方形邊長，使正方形可以蓋住目前已處理的點。把每一點 (x,y) 處理一遍:
   $d = max(d, min\{max(x - x_{min}, y - y_{min}), max(x_{max} - x, y_{max} - y)\})$
   其中 $max(x - x_{min}, y - y_{min})$ 等於左下正方形要蓋住 (x,y) 所需的最小邊長，與右上正方形所需的最小邊長比較，較小者用來更新 d。

   限制 space 的解法：以機率 $p = d(1/\varepsilon)logn/n$ (for some constant d) 取樣每個點，進行一般解法，得覆蓋範圍 R。
   space complexity 為 $O((1/\varepsilon)logn)$, time complexity 為 $O(n+(1/\varepsilon)logn)$.
   $Pr[|R \cap P| \geq (1 - \varepsilon)n]$
   $\geq 1 - Pr[|R \cap P| < (1 - \varepsilon)n]$
   $\geq 1 - O(n^3)(1/e^{\varepsilon pn})$
   $\geq 1 - (1/n^{\Omega(n)})$ when d is sufficiently large.