Homework 4: Data Streams

Đặng Mạnh Cường

1

Theo công thức Markov inequality:

$$P(X \geq a) \leq \frac{E(X)}{a}$$
$$\Rightarrow P_j = Pr[c_{j,h(i)_j} - F[i] \geq \epsilon t] \leq \frac{E[c_{j,h(i)_j} - F[i]]}{\epsilon t}$$

Mặt khác, theo tính chất:

$$E[c_{j,h(i)_j}] \leq F[i] + \frac{\epsilon}{e}(t - F[i])$$
$$\Rightarrow P_j \leq \frac{\frac{\epsilon}{e}(t - F[i])}{\epsilon t}$$
$$\Rightarrow P_j = Pr[c_{j,h(i)_j} \geq F[i] + \epsilon t] \leq \frac{t - F[i]}{et} \leq \frac{1}{e}$$

Xác suất để $c_{j,h(i)_j} \geq F[i] + \epsilon t, \forall j \in [1, log(\frac{1}{\delta})]$ là:

$$P = \prod_{j=1}^{log(\frac{1}{\delta})} P_j \leq \frac{1}{e}^{log(\frac{1}{\delta})} = \delta$$

Do $\tilde{F}[i] = Min(c_{j,h(i)_j}), \forall j \in [1, log(\frac{1}{\delta})]$

$$Pr[\tilde{F}[i] \geq F[i] + \epsilon t] = P \leq \delta$$
$$\Rightarrow Pr[\tilde{F}[i] \leq F[i] + \epsilon t] \geq 1 - \delta \text{ (DPCM)}$$