Lecture 20

Count-Min

q_i, $i = 1 \dots N$

$q \in I$ $\quad q_{i_1}, q_{i_2} \dots q_{i_w}$

Problem 1 – Point Query

Given i estimate q_i

Problem 2 – Inner Product (min of join)

q, L

$q \cdot b = \sum q_i b_i$

Problem 3 Range query

$b_i = \begin{cases} 1 & i \in R \\ 0 & \text{otw} \end{cases}$ $\quad \sum_{i \in R} q_i$

Stream model.

1° item q_i comes in arbitrary order

(i, v)

$(0, 13), (11, 12), (1, 2)$

2° Partial values of q_i on the stream

$(10, 13), (1, 12), (10, 1)$ → if all input tuples like $(10, \pm)$ $(1, \pm)$ set frequencies

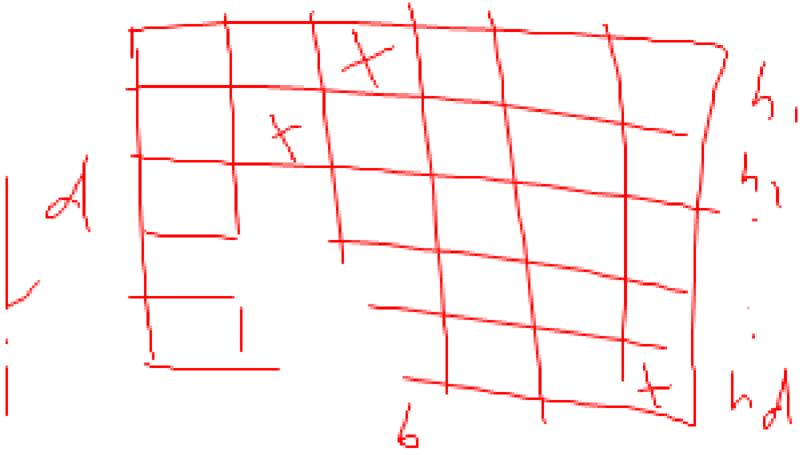
$(10, 14)$

AMS size of sketch $\sim \frac{1}{\varepsilon^2}$

$\varepsilon = 0.1 \quad \sim 100$

If sketch $\sim \frac{1}{\varepsilon} \quad \sim 10$

COUNT-MIN



(i, v_i)

For every $j = 1..d$

Compute $h_j(i) = k$

$A_{jk} \mathrel{+}= v_i$

Initially $A_{jk} = 0$

"Point Query" i in given

$\tilde{a}_i$

Fix j and k

$A_{jk} = \sum_i I_{i,j,k}\, a_i$

$I_{i,j,k} = \begin{cases} 1 & \text{if } x_i \text{ gets mapped to } k \\ 0 & \text{otherwise} \end{cases}$

$A_{jk} = \sum_{i:\, h_j(i)=k} a_i$

$\hat{a} = \min_{\substack{k = h_j(i) \\ j=1..d}} A_{jk}$