Lecture 11: Streaming Algorithm

Lecturer: Nisarg Shah   *Scribe: Lily Li*

11.1 Majority Algorithm

Recall our majority algorithm from last time: keep only one distinct element, add if you see the same element, subtract otherwise. Note: if a majority element does not exist this algorithm will not be able to detect this in one pass. You need another pass to make sure the element you picked is indeed a majority.

11.1.1 k-Heavy Hitters

We want to find k the elements $v_1, ..., v_k$ such that the number of occurrences of $v_i \geq \frac{n}{k}$. The algorithm is quite similar to the above. Now, we keep k different entires. We add the new entry if there is space among the k spots. If there is no space we *decrease all current counters* by one. The reason this works is by a charge like argument.

11.1.2 ϕ-Heavy Hitters

Let $0 \leq \phi \leq 1$. You must output all values that have frequency ϕm and all other returned values must have frequency at least $(\phi - \epsilon)m$ for some small ϵ. Note that the number of items depends on the input.

11.1.3 Sketching

(This idea is used in the Count-Min Algorithm) The idea of sketching is as follows:

11.2 Semi-streaming Models