Abstract

We consider the problem of approximating frequency moments in the streaming model. Given a stream $D = \{p_1, p_2, \ldots, p_m\}$ of numbers from $\{1, \ldots, n\}$, a frequency of $i$ is defined as $f_i = \{j: p_j = i\}$. The $k$-th frequency moment of $D$ is defined as $F_k = \sum_{i=1}^{\hat{n}} f_i^k$. In their celebrated paper, Alon, Matias, and Szegedy (STOC 1996) introduced the problem of computing a $(1 + \epsilon)$-approximation of $F_k$ with sublinear memory. We give upper bound of $O(n^{1-2/k})$ bits that matches, up to a constant factor, the lower bound of Woodruff and Zhang (STOC 12) for constant epsilon and $k > 3$ (joint work with Jonathan Katzman, Charles Seidell and Gregory Vorsanger, APPROX 14).

If time permits, we will present some recent results on universal sketches (joint works with Alan Roytman and Rafail Ostrovsky, RANDOM 15, and Stephen Chestnut, RANDOM 15), and $k$-median clustering on sliding windows (joint work Harry Lang, Keith Levin and Morteza Monemizadeh, SODA 16, FSTTCS 16).