Contribution ID: **44**

Type: **not specified**

Lower Bounds for Pseudo-Deterministic Counting in a Stream

Tuesday, July 11, 2023 3:30 PM (20 minutes)

Vladimir Braverman, Robert Krauthgamer, Aditya Krishnan and Shay Sapir

Abstract: Many streaming algorithms provide only a high-probability relative approximation. These two relaxations, of allowing approximation and randomization, seem necessary – for many streaming problems, both relaxations must be employed simultaneously, to avoid an exponentially larger (and often trivial) space complexity. A common drawback of these randomized approximate algorithms is that independent executions on the same input have different outputs, that depend on their random coins. \emph{Pseudo-deterministic} algorithms combat this issue, and for every input, they output with high probability the same "canonical" solution.

We consider perhaps the most basic problem in data streams, of counting the number of items in a stream of length at most n. Morris's counter [CACM, 1978] is a randomized approximation algorithm for this problem that uses $O(\log\log n)$ bits of space, for every fixed approximation factor (greater than 1). Goldwasser, Grossman, Mohanty and Woodruff [ITCS 2020] asked whether pseudo-deterministic approximation algorithms can match this space complexity. Our main result answers their question negatively, and shows that such algorithms must use $\Omega(\sqrt{\log n/\log\log n})$ bits of space.

Our approach is based on a problem that we call \emph{Shift Finding}, and may be of independent interest. In this problem, one has query access to a shifted version of a known string $F \in \{0,1\}^{3n}$, which is guaranteed to start with n zeros and end with n ones, and the goal is to find the unknown shift using a small number of queries. We provide for this problem an algorithm that uses $O(\sqrt{n})$ queries. It remains open whether $poly(\log n)$ queries suffice; if true, then our techniques immediately imply a nearly-tight $\Omega(\log n/\log\log n)$ space bound for pseudo-deterministic approximate counting.

Presenter: SAPIR, Shay

Session Classification: Track A-4