

Ganguly, Sumit

Lower bounds on frequency estimation of data streams (extended abstract). (English)

Zbl 1142.68331

Hirsch, Edward A. (ed.) et al., Computer science – theory and applications. Third international computer science symposium in Russia, CSR 2008 Moscow, Russia, June 7–12, 2008. Proceedings. Berlin: Springer (ISBN 978-3-540-79708-1/pbk). Lecture Notes in Computer Science 5010, 204-215 (2008).

Summary: We consider a basic problem in the general data streaming model, namely, to estimate a vector $f \in \mathbb{Z}^n$ that is arbitrarily updated (i.e., incremented or decremented) coordinate-wise. The estimate $\hat{f} \in \mathbb{Z}^n$ must satisfy $\|\hat{f} - f\|_\infty \leq \epsilon \|f\|_1$, that is, $\forall i (|\hat{f}_i - f_i| \leq \epsilon \|f\|_1)$. It is known to have $\tilde{O}(\epsilon^{-1})$ randomized space upper bound, $\Omega(\epsilon^{-1} \log(\epsilon n))$ space lower bound and deterministic space upper bound of $\tilde{\Omega}(\epsilon^{-2})$ bits. We show that any deterministic algorithm for this problem requires space $\Omega(\epsilon^{-2}(\log \|f\|_1)(\log n)(\log^{-1}(\epsilon^{-1})))$ bits.

For the entire collection see [Zbl 1136.68005].

MSC:

- 68P05 Data structures
- 68Q25 Analysis of algorithms and problem complexity
- 68Q45 Formal languages and automata

Cited in 2 Documents

Full Text: [DOI](#)