Sublinear Time and Space Algorithms 2016B – Problem Set 4

Robert Krauthgamer

Due: June 19, 2016 (corrected version)

General instructions: Please keep your answers short and easy to read. You can use results, calculations or notation seen in class without repeating them, unless asked explicitly to redo them.

1. We saw in class an algorithm that approximates the average degree d in a graph on n vertices within factor $(2+\varepsilon)$ in time $O((\frac{1}{\varepsilon})^{O(1)}\sqrt{n})$.

   Explain how to improve the runtime when the algorithm is given $1 \leq d_0 \leq d$, i.e., a lower bound on the average degree.

2. Show that every streaming algorithm, even a randomized one, for computing ℓ_1 exactly in $\mathbb{R}^n$, requires storage of $\Omega(n)$ bits.

   Hint: You obviously must use a stream with deletions.

3. Show that every algorithm, even a randomized one, for $(1+\varepsilon)$-approximating the ℓ_0-norm of an *insertions-only stream* in $\mathbb{R}^n$ for $2/\sqrt{n} \leq \varepsilon < 1$, requires storage of $\Omega(1/\varepsilon^2)$ bits.

   (In the proof seen in class, the input is of the form $x - y$ and thus contains deletions.)

   Hint: Observe that $2\|x+y\|_0 = \|x\|_0 + \|y\|_0 + \|x-y\|_0$ for all $x, y \in \{0,1\}^n$.

Extra credit:

4. Prove the every algorithm that returns the exact median of a stream of odd length n, requires storage of $\Omega(n)$ bits, even if the algorithm is randomized and items come from a domain of size $O(n)$.

   Hint: use reduction from Indexing.