Sublinear Time and Space Algorithms 2018B – Problem Set 4

Robert Krauthgamer

Due: June 10, 2018

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. Let $y \in \mathbb{R}^n$ be the frequency vector of an input stream in the turnstile model (i.e., allowing insertions and deletions), and suppose its coordinates are integers in the range $[-n^2, n^2]$.

   Design a linear sketch that detects whether $|\operatorname{supp}(y)| = 1$ using storage requirement of $O(1)$ words (i.e., $O(\log n)$ bits), not counting storage for algorithm's random coins. Its success probability should be at least $1 - 1/n$.

   Hint: Use a variant of the AMS sketch (tug-of-war) with large random coefficients.

2. Analyze Algorithm 2 below for counting triangles in a graph given as a stream, and show that with constant high probability, the additive error is $|\tilde{T} - T| \leq \varepsilon t$. Can this algorithm be applied also for dynamic graphs (i.e., a stream of edge insertions and deletions)? Explain how/why not.

   Notation (similar to class): Assume $t > 0$ is a known lower bound for the actual number of triangles T, and let x_S count the number of edges internal to the vertices $S \subset V$.

   Algorithm 2

   1. Init: pick $k = O(\frac{n^3}{\varepsilon^2 t})$ random subsets $S_1, \ldots, S_k \subset V$ each of size 3 (with replacement)
   2. Update: maintain $x_{S_1}, \ldots, x_{S_k}$ (explicitly)
   3. Output: compute $z = \sum_{i \in [k]} \mathbb{1}_{\{x_{S_i} = 3\}}$ and $N = \binom{n}{3}$, and report $\tilde{T} = \frac{N}{k} \cdot z$

   Hint: Use Chebyshev's inequality.

   Extra credit:

3. Show how to improve Algorithm 2 above by choosing the random sets S_i only from those sets S that satisfy $x_S \geq 1$. The resulting **Algorithm 2'** should have space requirement $k' = O(\frac{mn}{\varepsilon^2 t})$ words, and work also for dynamic graphs.

   Hint: Use ℓ_0-samplers and an estimator for $N' = \|x\|_0$.