Abstract:
CountSketch is a popular dimensionality reduction technique that maps vectors to a lower dimension using randomized linear measurements while supporting the recovery of the L2-heavy hitters of the input vectors. We study the robustness of CountSketch in adaptive settings where input vectors may depend on the output from prior inputs. Adaptive settings arise in processes with feedback or with adversarial attacks. We show that
(1) CountSketch itself is not robust, and can be attacked with a number of queries of the order of the sketch size.
(2) A slight variation of CountSketch is robust, allowing for quadratic number of queries in the sketch size, which is an improvement factor of $\sqrt{k}$ over prior work (for $k$ heavy hitters).
Based on a joint work with Edith Cohen, Xin Lyu, Jelani Nelson, Tamas Sarlos, and Moshe Shechner.