UNIVERSITÀ DI PISA

Eight hands-on: Count-min sketch: range queries

Algorithm Design (2021/2022)

Gabriele Pappalardo
Email: g.pappalardo4@studenti.unipi.it
Department of Computer Science

April 2022

1 Introduction

Consider the counters $F[i]$ for $1 \leq i \leq n$, where n is the number of items in the stream of any length. At any time, we know that $||F||$ is the total number of items (with repetitions) seen so far, where each $F[i]$ contains how many times the item i has been so far. We saw that CM-sketches provide a FPTAS $F'[i]$ such that $F[i] \leq F'[i] \leq F[i] + \epsilon||F||$, where the latter inequality holds with probability at least $1 - \delta$.

2 Solution

Before explaining the solution to the problem, we introduce the *Heavy Hitters* problem and how the *Count-Min Sketch* data structure solves it.

2.1 Heavy Hitters

What is the Heavy Hitters problem?

2.2 Count-Min Sketch

Describe the Count-Min Sketch solution and how does it work.