

PRESENTATION OUTLINE: Parallel and Fast Algorithms for Finding Top-K Frequent Elements

Mostafa Hosseini
School of Computer Science
Carleton University
Ottawa, Canada K1S 5B6
MostafaHosseini@cmail.carleton.ca

November 24, 2022

1 Contents

- 1 slide - 1 minute
- Contents of presentation
- Different sections

2 Introduction

- 1 slide- 2 minutes
- What is the subject (finding top-k frequent items in the massive data stream)
- Why this topic is important
- What are the applications of the subject (examples)
- What parameters are noticable in this subject

3 Solution

- 1 slides- 2 minutes
- Exact algorithm for solving the problem
- Approximate algorithm for solving the problem
- Overview of different types of approximate algorithm(Sketch based and counter based)
- Advantages and disadvantages of sketch based and counter based algorithm

4 Count Min Sketch

- 2 slides- 2 minutes
- What is Count min Sketch(CMS)
- How does the CMS work
- An example about CMS
- Computation time of CMS

5 Asketch

- 2 slides- 2 minutes
- Explaining the Asketch method
- An example about Asketch

6 Topkapi

- 2 slides- 2 minutes
- Explaining the Topkapi method
- An example about Topkapi

7 Proposed method

- 1 slide- 1 minute
- Explaining the proposed method
- What is the difference between proposed method and other methods

8 Parallelizing the methods

- 1 slide- 2 minutes
- Multi-Core Parallelism
- Describing how to make the parallel implemetation

9 SIMD Intrinsic

- 1 slide- 2 minutes
- How to implement SIMD sort and finding minimum item for the filter in one core
- Advantages of using SIMD Intrinsic
- Overview of SIMD Intrinsic code for 128 bits and 256 bits vector

10 Implementation setup

- 1 slide- 1 minute
- Describing the language, processors, dataset, etc. used for the project

11 Result

- 2 slides- 2 minutes
- Comparing different methods
- Comparison of different $K = 32, 64, 128, 256$
- Speedup comparison for proposed method

12 Conclusion

- 1 slide- 1 minute
- What did we do in this project
- What did we get from this project
- What can be as a future work

13 Questions

- 1 slide
- Question1
- Question2
- Question3