
Parallel Data Mining

Jason Ma & Kanishk Tantia
Introduction to High Performance and Parallel Computing
Harvey Mudd College
jyma@g.hmc.edu & ktantia@g.hmc.edu

Abstract

- 1 Introduction**
- 2 Data Processing**
- 3 Algorithms Tested**
 - 3.1 FastCDC**
 - 3.2 RabinCDC**
 - 3.3 AECDC**
- 4 Experimental Approach**
 - 4.1 Practical Traces**
 - 4.2 Synthetic Traces**
- 5 Results**
- 6 Conclusions**
- 7 References**