Real-Time Large-Scale Data Analytics and Information Retrieval in Practice

Aleksandar Bradic, Igor Bogicevic

2009

Contents

Chapter 1

Introduction

1.1 Enter the real time

1.2 Fundamental issues

Chapter 2

The nature of large-scale data

Chapter 3

The challenges of real-time information processing

3.1 Problem description

Chapter 4

Fundamental Algorithms in Data Analytics and IR

4.1 Statistical analysis framework

4.1.1 Regression analysis

4.1.2 Forecasting

4.1.3 Parameter estimation

Chapter 5

Advanced Algorithms

5.1 Online learning algorithms

5.2 Kernel Methods

Chapter 6

Software toolkits for large-scale data analysis

6.1 Hadoop

6.2 Mahout

Chapter 7

Large-scale IR Cookbook

Chapter 8

Moving from batch to real-time

Chapter 9

Real-world real-time applications

9.1 Web Analytics

9.2 Media analysis

9.3 Finance

9.4 Online collaboration

Chapter 10

Algorithms and Data Structure in support of large-scale real-time framework

10.1 Randomized Algorithms

10.2 Queue-based structures

Chapter 11

VoidBase : queue-based computing framework

11.1 Overview

11.2 Cookbook