IT UNIVERSITY OF COPENHAGEN

Clustering Player Behaviors Efficiently with K-Means in Map-Reduce

by

Sigurdur Karl Magnusson

A thesis submitted in partial fulfillment for the
degree of Master of Science in IT

in the

Computer Science
Software Development and Technology

Supervisors
Julian Togelius, Rasmus Pagh

March 2013

"Write a funny quote here."

If the quote is taken from someone, their name goes here

Abstract

The abstract is a short summary of the thesis. It announces in a brief and concise way the scientific goals, methods, and most important results. The chapter "conclusions" is not equivalent to the abstract! Nevertheless, the abstract may contain concluding remarks. The abstract should not be discursive. Hence, it cannot summarize all aspects of the thesis in very detail. Nothing should appear in an abstract that is not also covered in the body of the thesis itself. Hence, the abstract should be the last part of the thesis to be compiled by the author.

A good abstract has the following properties: *Comprehensive:* All major parts of the main text must also appear in the abstract. *Precise:* Results, interpretations, and opinions must not differ from the ones in the main text. Avoid even subtle shifts in emphasis. *Objective:* It may contain evaluative components, but it must not seem judgemental, even if the thesis topic raises controversial issues. *Concise:* It should only contain the most important results. It should not exceed 300–500 words or about one page. *Intelligible:* It should only contain widely-used terms. It should not contain equations and citations. Try to avoid symbols and acronyms (or at least explain them). *Informative:* The reader should be able to quickly evaluate, whether or not the thesis is relevant for his/her work.

An Example: The objective was to determine whether ... (*question/goal*). For this purpose, ... was ... (*methodology*). It was found that ... (*results*). The results demonstrate that ... (*answer*).

Acknowledgements

The acknowledgements and the people to thank go here, don't forget to include your project advisor. . .

Contents

List of Figures

List of Tables

Abbreviations

LAH   **L**ist **A**bbreviations **H**ere

Chapter 1

Chapter Title Here

1.1 Motivation

1.2 Problem Statement

1.3 Contributions

1.4 Methodology

1.5 Project Outline

Chapter 2

Background Theory

2.1 Player Behaviour

2.1.1 Game Metric

2.1.2 Features

2.2 Clustering

2.2.1 K-Means

2.2.2 Streaming and Data Streams

2.3 Map-Reduce and Hadoop

Chapter 3

Related Work

3.1   Player Behaviours

3.2   Clustering

3.2.1   Streaming

3.2.2   Data Streams

Chapter 4

Methodology

4.1 Game Events Data

4.1.1 Building Features

4.2 K-Means in Map-Reduce

4.2.1 Euclidean distance

4.2.2 Normalized Histograms

4.3 Experimental Set-Up

Chapter 5

Results and Discussion

5.1   Results

5.2   Discussion

Chapter 6

Conclusions

Appendix A

Appendix Title Here

Write your Appendix content here.