Cloud Programming: Group 06 Final Project Proposal

Ming-Chang Chiu 100060007
Mack Teng 100060025

April 27, 2015

1 Introduction

In this project we are going to develop a recommender system using MapReduce framework. A research group from Stanford University provides millions of data from amazon, which are clearly categorized and consist of important informations such as prices, ratings, reviews, etc.. We aims to mine this massive dataset and discern if there are follow-the-leader effects with those reviews. We are going to find out whether follow-the-leader effect exists with respect to category or to product, and if the effect is conspicuous then we can tell amazon the improve the recommender system or for future market use.

2 System Details

We will develop two layers of analysis. First, the judgement of reviews. The system should be able to discern a reviewer's attitude toward a commodity based on his rating and text review. Second, the review pattern. We ought to discover whether certain category or certain kind of product would have follow-the-leader effect.

3 Components of the system

Algorithm: We will first find or develop a NLP(Natural Language Processing) algorithm which can tell us what is the stance of a certain review about a product and algorithms that can find patterns of reviews.

Platform: Our work will be constructed on AWS using S3, ELB, and so on.

Web Interface: Eventually, we will write a simple interface giving users the right to arbitrarily create a product which can be categorized and then see the result we predict.

4 Expected Results

Input: Users can create a product, put it in any category of amazon and write a review and set a rating manually.

Output: The system should tell us whether our review is a positive or negative one, and whether this review will affect later customers' opinion about it.