Visual Detection of Singularities in Recommendation Systems

A. Colantonio¹, R. Di Pietro², M. Petrocchi³, A. Spognardi³
¹Bay31 AG, Switzerland ²Bell Labs Alcatel Lucent, Paris, France ³CNR IIT, Pisa, Italy

Introduction

Reviews can largely influence purchase decisions: if most reviews are positive, people are inclined to buy the product; on the contrary, negative reviews will persuade the user to look for alternatives. This fact leads to the submission of fake reviews with the main intent of twisting product perception [2]. Academia has shown a growing interest towards recommendation systems and a flourishing literature exists on deceptive reviews [4].

We propose analyzing reviews resorting to visual analytics—the science of analytic reasoning facilitated by visual interfaces [3]. This strengthens human interaction with complex graphical displays, putting humans in the loop allows for a better correlation of knowledge with review data. Moreover, visualization enables the analyst to visually mine data by means of human reasoning.

We propose a visual approach for singularity detection in review platforms. Our idea is to complement existing review assessment techniques by providing a "visual workbench" for review analysts. We adapted results coming from IT security [1] to implement a novel methodology for review analysis.

Methods

The problem that we address is offering a graphical way to effectively navigate existing reviews in a system, showing at a glance what it would take a lot of data, processing and time to expound. Among all possible graphical representations, we focus on matrix-based visualization, a two-dimensional graphical representation of a matrix, where rows, columns, and cells have assigned a specific meaning. The main problem to address when depicting a matrix is identifying a proper sorting of rows and columns: it can be proven that rearranging rows and columns allows to highlight the main patterns embedded in the represented data. To identify the best row/column sorting, we use a fast algorithm referred to as ADVISER (Access Data VISualization) [1], which is able to provide a compact representation of patterns embedded in a binary matrix.

Our approach starts from specifying how to construct a matrix-based visualization out of review data. To this aim, we extracted real data from TripAdvisor (http://www.tripadvisor.com) and Booking.com (http://www.booking.com) websites. The extracted dataset consists of a list of reviews, each one made up of: user ID, hotel name, review date, review text, rating. We discuss the matrix representation obtained by assigning each row to review weeks, sorted from top to bottom in reverse chronological order; columns represent hotels, sorted according to the above mentioned ADVISER algorithm; finally, a cell is "filled" when there is at least one review for the given hotel in the given week.

Results

Tripadvisor reviews removed after September 10, 2013

Shared reviews between TripAdvisor and Booking.com

Reviews in TripAdvisor for hotels in common between the two platforms

Reviews in Booking.com for hotels in common between the two platforms

Reviews in TripAdvisor and not in Booking.com

Reviews in Booking.com and not in TripAdvisor

Conclusions

We have highlighted how to interpret raw review data as input to binary matrices that are the basis for the visualization and we have showed the effectiveness and quality of the proposed solution via an extensive experimental campaign over real data extracted from TripAdvisor and Booking.com. While the proposed approach is not a definitive answer to the vexed issue of detecting misleading reviews, we believe that our novel approach paves the way for further contributions in the area of visual analytics applied to recommendation systems.

Bibliography


Acknowledgments

This research has been partially supported by the MIB (My Information Bubble) project of Registo.it