Analysis and Evaluation of a Framework for Sampling Database in Recommenders

Hodjat Hamidi, Department of Information Technology, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran
Reza Mousavi, Department of Information Technology, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

ABSTRACT

In this paper the authors proposed a database sampling framework that aims to minimize the time necessary to produce a sample database. They argue that the performance of current relational database sampling techniques that maintain the data integrity of the sample database is low and a faster strategy needs to be devised. The sampling method targets the production environment of a system under development that generally consists of large amounts of data computationally costly to analyze. The results have been improved due to the fact that the authors have selected the users that they had more information about them and they have made the data table denser. Therefore, by increasing the data and making the rating more comprehensive for all the users they can help to produce the more and better association rules. The obtained results were not that much suitable for Jester dataset but with their proposed methods the authors have tried to improve the quantity and quality of the rules. These results indicate that the effectiveness of the system greatly depends on the input data and the applied dataset. In addition, if the user rates more number of the items the system efficiency will be more increased.

KEYWORDS

Association Rules, Database Sampling Framework, Recommender Systems, Sampling

1. INTRODUCTION

Recommender systems are an important facet of Internet-based sellers’ advertising and promotion strategies due to their ability to deliver shopping advice, stimulate consumers’ purchase desires and boost sales (Jannach et al., 2011). However, when managers introduce recommendation services, do they need to take into account that men and woman may react differently to the advice? This is important because advertising is the major way in which marketers communicate with different target segments in the traditional market, and recommendations have played a similar role in the e-commerce market. Moreover, for many years, gender has been considered to be the most useful basis for market segmentation because both segments are profitable and easy to target (Adomavicius et al., 2005). Consequently, managers must understand whether there are decision-processing differences between men and women in order to produce effective recommender system advice for each segment. Since that start, the field has advanced through both basic research and commercial development to the point where today recommender systems are embedded in a wide range of commerce and content applications (both online and offline), where recommender systems handbooks and texts have been published (e.g., Ricci et al., 2011; Wu et al., 2016; Kumar, 2016; Ye et al., 2016; Hamidi et al. 2017),

DOI: 10.4018/JGIM.2018010103

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
where universities are offering courses on recommender systems, and where there is a dedicated annual conference on the topic. The scope of recommender systems has also broadened; while the term originally grew out of work in collaborative filtering, it quickly expanded to include a broader range of content-based and knowledge-based approaches. While such systems are important, we limit our focus to recommender systems that are based on collaborative filtering, though many of the interface issues we discuss apply to recommenders based on different approaches. This limitation reflects both our own expertise and the practical limitations of addressing so broad a field in a single article. We do not attempt to offer a comprehensive review of past algorithmic research.

Indeed, there have been a number of thorough surveys that focus on the algorithms behind recommenders (Burke 2002; Ekstrand et al., 2011; ; Hamidi et al., 2017; Herlocker et al., 2004), and we refer the interested reader to them. Rather, we present an overview of the most important developments in the field that touch on the user experience of the recommender. By user experience we mean the delivery of the recommendations to the user and the interaction of the user with those recommendations. The user experience necessarily includes algorithms, often extended from their original form, but these algorithms are now embedded in the context of the application. Our review looks at research grounded in specific recommender systems and their evaluations, and stands in contrast to Knijnenburg et al. (2012) which approaches user experience from more of an experience-model and social-experimental approach. In the rest of this section we highlight the main directions of work in the early years of recommender systems, including the beginning of the shift away from thinking of recommenders as prediction engines to considering them in the context of user experience.

In this paper, we propose a very fast sampling method that maintains the referential integrity of the sample database intact. The sampling method targets the production environment of a system under development, that generally consists of large amounts of data computationally costly to analyze.

The remaining part of the paper is organized as follows: The literature review of the subject is presented in the second section. In the third section, the user-recommender lifecycle will be introduced. The very fast database sampling and result and evaluation will be presented in the fourth and fifth section. In the sixth sections of the article, the conclusion will be provided.

2. RELATED WORK

As recommender use expanded rapidly among online retailers and online content providers, applications of recommenders grew more diverse but the underlying algorithms converged to a few particularly useful ones. The classic recommender algorithm described above, known as user-user collaborative filtering, because the correlation is measured between pairs of users, was widely recognized as providing high-quality predictions and recommendations (see, for instance, (Breese et al., 1998)), but in practice often performed too slowly to be suitable for real-time use in applications with hundreds of thousands or millions of users. Item-item collaborative filtering (Sarwar et al., 2001) was developed as an alternative algorithm; it builds correlations between pairs of items, and then computes recommendations by finding items with high similarity to the set of items already rated favorably by the user. Many ecommerce stores have many more customers than items, and more stable relationships between items than between customers. In these stores the item–item algorithm has faster online response time than the user–user algorithm, especially if the item relationships are precomputed. The item–item algorithm, which also extends nicely to unary rating sets (sets where the database has either positive information or no information at all, such as sales data), quickly became popular in commercial applications.

Alternative algorithms based on dimensionality reduction (Billings et al., 1998, Mousavi et al., 2016; Bimonte et al., 2017; Esposito et al., 2016) showed early promise for commercial application and have been adapted in many ways to deliver high performance and high-quality recommendations. These methods, commonly based on singular value decomposition, start with the recognition that a user–item ratings matrix actually has too many independent dimensions and thus loses some of
Related Content

Metastructuring for Standards: How Organizations Respond to the Multiplicity of Standards
Ronny Gey and Andrea Fried (2018). Corporate and Global Standardization Initiatives in Contemporary Society (pp. 252-276).
www.igi-global.com/chapter/metastructuring-for-standards/197468?camid=4v1a

Open Source Software Development Process Model: A Grounded Theory Approach
Keng Siau and Yuhong Tian (2013). Journal of Global Information Management (pp. 103-120).
www.igi-global.com/article/open-source-software-development-process-model/99667?camid=4v1a
Complex Network Perspective on Collaboration in ICT Standardization
www.igi-global.com/chapter/complex-network-perspective-on-collaboration-in-ict-standardization/197458?camid=4v1a

Analysis and Evaluation of a Framework for Sampling Database in Recommenders
www.igi-global.com/article/analysis-and-evaluation-of-a-framework-for-sampling-database-in-recommenders/190490?camid=4v1a