A Survey of the Application of Association Rules in Library Services

Chu-feng ZHOU¹, Xin-xin GUAN¹, Jun-wei HOU³, Yun-fei ZI¹, Ye-li LI¹ and Qing-tao ZENG¹,²,³

¹School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China
²Tsinghua University, Computer Science and Technology Postdoctoral Research Station, Beijing 100084, China
³System Engineering Research Institute, CSSC, Beijing 100094, China
*Corresponding author

Keywords: Data mining, Association rules, Libraries, Applications, Overview.

Abstract. The article analyzes the application of association rules at libraries both at home and abroad on the application of related documents in libraries, and analyzes them from the application of association rules in digital library personalized recommendation services, library resource configuration management and reader behavior analysis. The status of application, and put forward the prospects for its future development.

Introduction

Data mining generally refers to the process of searching for information hidden in an algorithm from a large amount of data. Extract previously unknown, valid, and operational information from large data warehouses or databases, and then use this derived information for critical business and strategic decisions.

With the development of social informatization, data mining technology has been applied in a wide range of fields in society. Through the analysis, processing, and application of data, users can realize customized and personalized services, and effectively predict and grasp various types of needs. Deep processing applications. This has provided a new direction for the development of the library. The library has accumulated a large amount of continuously updated data in the process of informatization construction, and applied research on the direction of the library through association algorithms to improve the book recommendation, reader analysis, and book planning. And other library services.

Algorithm Introduction

In data mining, Apriori algorithm is one of the most influential algorithms for mining frequent itemsets of Boolean association rules. Its core is a recursive algorithm based on the idea of two-stage frequency set.

The basic idea is to scan the entire database first to form an initial frequent "1-item set" set L1. After frequent pruning of the "1-item set" through the predefined support and confidence, frequent formation of L1 through pruning occurs. The “2-item set” L2 is set, then L2 is used to find L3, and iteration is continued until a frequent “k-item set” cannot be found, and the desired association rule [1] is generated with the minimum support and the minimum confidence as criteria.

The traditional Apriori algorithm mainly has two defects:
- 1. The set of candidate items that may be generated is huge resulting in a combined explosion.
- 2. After generating a set of candidate items, the Apriori algorithm must scan the database to determine if each item is a frequent item.
Therefore, in order to solve these problems, many improvements have been proposed: the correlation analysis algorithm proposed by Han Jiawei et al. in 2000, and an FP-Growth algorithm called Frequent Pattern Tree data structure is used in the algorithm [2], based on Hash tree DHP algorithm, AprioriTid algorithm and Partition algorithm [4], by S Shirgaonkar improved Apriori algorithm [3].

**Foreign Research Status**

As early as in 1999, the LITA (Library and Information Technology Association) experts from the Library and Information Technology Group of the American Library Association (ALA) affiliated with the American Library Association put forward seven trends in library technology development at a seminar. . The United States, Canada and other countries have successively developed technologies for library applications. For example: the earlier emergence of the Mylibrary system. Professor Michael Cooper of the University of California, USA, made a scientific prediction of user behavior through the law of stay of different types of users. The Arrowsmith system developed by the University of Chicago's Swansan used in-depth excavation of database literature information to obtain intrinsic links between documents. [10][13]

**Research State in China**

Domestic start is relatively late compared to foreign countries, but there are still some nice and relatively good results. The China Knowledge Network project initiated by Tsinghua University and Tsinghua Tongfang has launched personal digital libraries, Renmin University of China libraries, and Shenzhen University books. Library and other personalized library recommendation system. However, as a whole, there is a problem that the system functions are relatively simple, the service content is in a few forms, and the degree of service initiative is low. Therefore, the information cannot be effectively pushed, and the level of interaction between the administrator and the user is low. [12]

**Research Hotspot**

**Reader Behavior Analysis**

Through the questionnaire survey, the association rules mining with library data, and analyzing the readers’ borrowing methods, borrowing quantity, and borrowing reasons in different regions and different ages form reliable association rules to guide the orientation of the library to better provide for different groups. The service, in order to change the library's service capabilities, to consolidate and attract more readership. At the same time, it can analyze the readers who have bad borrowing behaviors, and seek a relative solution from the perspective of coping. [5][8]

**Library Resource Allocation**

Compared to the purposeless purchase of books and management of books, the library hopes to obtain a reasonable solution through effective analysis. Through the library readers basic information, borrowing bibliographic attributes and book circulation data to achieve association rules mining, looking for Reliable association rules among various types of books make it possible for libraries to ensure reasonable distribution of various types of books and materials under limited acquisition costs. At the same time, the internal library of the library can be reasonably racked up to achieve clear and easy management. [6]

**Personalized Book Recommendation**

In a complex network environment, users need precise personalized services, which are also referred to as customized services. Personalized services through the library to borrow data to achieve association rules mining, according to the explicit needs of readers, the reader's age, education, personality and other information analysis, to find the hidden rules, and the mining rules applied to
the personality. In the recommendation system, it automatically searches the collection database, actively guides the borrowing behavior of the readers, and accurately recommends the books they need. [7][9][10]

**Mobile Library**

With the rapid development of mobile network technology today, the rise and popularity of mobile devices have made the service approach convenient, practical and timeless. Obtaining corresponding services through mobile phone devices is more in line with the needs of young people today, and a novel library service model—library mobile service model has emerged. The mobile library accesses library resources through mobile devices such as smartphones, and accesses functions such as book retrieval, borrowing, and business inquiries. Because the mobile library's most obvious feature is its mobility, it also has the characteristics of real-time and initiative, so that the library can obtain more user-active, real-time data to analyze the user's association rules. [11]

**Prospects for Future Development**

Library association rules mining is an emerging technology field. The My Library-Christ church College of Education at the Christchurch School of Education in New Zealand and the My Gateway-University of Washington Libraries at the University of Washington Library are already in practice. Proved its role. This article summarizes the current research hotspots based on association rules reader behavior analysis, personalized recommendation, resource allocation and mobile library application direction, but we can still see that a large number of domestic studies are in the literature theoretical research, practical application is less, and follow The continuous development of various Internet technologies will also be combined with such technologies as "cloud computing" and mobile Internet technologies to achieve a more complete design of the library system.

**Acknowledgment**

Thanks to Beijing science and technology innovation service capability construction project (PXM2017_014223_000055), Beijing Institute of Graphic Communication faculty construction -- doctoral start-up project (27170118003/007) and Beijing Institute of Graphic Communication 2018 R&D Project (Ec201802), Publication data asset appraisal platform construction (04190118002/039).

**References**


[8] Hongyun Li, the application of association rule in library [D], Shanghai: East China Normal University. 2009.


[11] Chao, Huang Research and implementation of mobile library based on Data Mining [D], Xi’an: Xi’an University of Architecture and Technology, 2015.


[18] Li Yan, Lv Peng and Li Long, Research on personalized services of University Library Based on the Big Data Ming and Decision Analysis System [A], Library, Document & communication 2016 No. 2. pp. 60-68.


[22] Bing Wang, Applied research on Data Ming in Library services of High school [D], Xi’an: Xi’an University of Technology, 2010.


