

Data Mining Applications in Library and Information Services

Dr. T. Anuradha^{#1}

[#] Associate Professor, Dept. of Computer Science, Dravidian University, Kuppam, A.P, India -517426,

Abstract— Schools, Colleges and Universities; local, state and central governments as well as non-government organizations and large commercial organizations maintain their own libraries. Computerized library systems store and maintain huge volumes of data about their collections, user information, circulation information, resource information, and possibly bibliographic search logs.

The increased use of information and communication technologies, database technologies, the Internet and World Wide Web has lead to the development of digital libraries. These digital libraries are serving a huge variety of different users and playing important roles as repositories and providers of Information and knowledge. Data mining techniques can be applied to the huge data related to the library systems to improve the quality of services provided by libraries and to improve the effectiveness of internal functional process of libraries. This paper presents applications of data mining in digital libraries.

Index Terms— Digital libraries, Bibliomining, Autonomous Citation Indexing, Applications of Data Mining,

I. INTRODUCTION

The latest database technologies, the Internet and World Wide Web, Information and Communications technologies had led to the development of digital libraries. These digital libraries are providing information and knowledge to a huge variety of different users.

Knowledge Discovery in Databases (KDD) is the process of extracting unknown, very valuable and hidden information from huge databases. KDD involves extraction, cleaning, storing, data mining, visualization and interpretation. Data Mining is a step of applying algorithms or techniques in the process of KDD. Data mining can be applied to huge volumes of library collections data, users and circulation information, bibliographic data etc. to improve the library services quality to effectively implement internal functional process of libraries.

II. RELATED WORK

The application of Information technology in library systems has lead to the establishment of Digital Libraries by integrating Library systems, database development, network technologies, online search and retrieval etc.

In [RYKDY 2008], data mining was integrated with Bayesian Networks to develop a personalized

recommendation into library services. Association rules were generated from readers' book interest data and readers were grouped based on their interest. Association rules were also generated between books from the book borrowing history. Based on these association rules Bayesian network structure model was constructed. This system recommends a user the Top-N list of other related books suitable to the reader.

Huge amounts of information have been available in libraries for long time. The revolution the web has brought to information dissemination is not much due to the availability of huge data, but because of the improved efficient accessing of that data. In [LGB1999], authors used Autonomous Citation Indexing (ACI) to evaluate the importance of individual contributions more quickly and accurately and showed that by incorporating ACI in digital libraries, scientific literature can be organized effectively and the efficiency of dissemination of information and feedback can be improved.

The principal purpose of the decision support system for libraries [BZCBZ1999] is to provide information regarding books usage pattern, periodicals and electronic services. The authors of [BZCBZ1999] proposed a prototype for a decision support system for the Kansas State University libraries and presented a solution to data integration incompatibilities which are due to inclusion of external data. This integrated library decision support system showed variation of library use patterns among different disciplines which can be used by library administration for adjusting and/or justifying purchases and licenses.

The authors [EA2011], assumed that researchers' borrowing history is an important source of information. A methodology consisting of four main steps is proposed based on this assumption. These four steps are: creation of data warehouse based on library operational data, creation of knowledge map for library subjects and their usage, analyzing knowledge map to choose subjects, and finally, from knowledge map analysis applying data mining techniques to find active researchers in selected subjects. Clustering is applied to researchers based on their research interest which helps decision makers regarding research collaborations. A knowledge map as a visual representation of usage trends of academic library is presented to describe virtual interest groups based on item use information.

In [Z2007], authors introduced the concepts of personalised services in digital library, data mining techniques and their functions; described the process of constructing the personalized services system based on data mining.

In [M2007], data mining technology is introduced and applications of data mining technologies in digital libraries are discussed. A model based on web-based data mining technology is analyzed and each constituent function of the model is illustrated.

[PPU2004] provided an overview of applications of Data Mining in Library and Information Services. Bibliomining refers to the applications of data mining techniques to explore patterns in bibliographic information. Bibliomining provides an insight into the learning and knowledge needs of library users, and therefore customize the library services in a better way to meet these needs. Bibliomining can also be used to predict the demand of new items and to reduce inventory losses. Bibliomining can be used to predict the future needs of the users.

[CR2004] authors proposed a method to develop Digital Libraries using data warehousing and data mining techniques. Architecture for the development of digital libraries, based on the Data warehousing approaches is presented. Data warehousing components are mapped to Digital librarying components and data warehousing process is mapped to digital librarying process. Applications of some data mining techniques such as web content mining and web usage mining in the three areas of digital librarying processes such as extraction, transformation and OLAP are proposed. These techniques allow the filtering of the really interesting documents in the extraction process, Dublin Core metadata generation using the key words found in the selected documents, and automated refresh of the library and revealing the needs of users through the analysis of their accesses standards.

In the paper [M2011], authors described basic concepts of data mining, web mining, and text data mining; major steps of data mining and functions of data mining such as association analysis, clustering, concept description, time-series pattern, deviation detection. Structure mining, content mining and usage mining are also described in the context of applications of data mining in digital library. This paper also described application meaning of data mining in the context of digital library. Finally it analyzes the key problems of implementation of data mining in digital libraries and takes a case study on the analysis of low utilization readers in Linyi university library using decision tree technique.

[MMG2004] listed some benefits of data mining in digital libraries like applications of data mining techniques for information retrieval for better browsing and searching; to discover unidentified knowledge areas; for text analysis tasks such as discipline extractions etc.

III. CONCLUSIONS:

Schools, Colleges and Universities; local, state and central governments as well as non-government organizations and large commercial organizations maintain their own libraries. Computerized library systems store and maintain huge volumes of data about their collections, user information, circulation information, resource information, and possibly bibliographic search logs. These digital libraries are serving a huge variety of different users and playing important roles as repositories and providers of Information and knowledge.

Data mining techniques can be applied to the huge data related to the library systems to improve the quality of services provided by libraries and to improve the effectiveness of internal functional process of libraries. This paper presented some applications of data mining in digital libraries.

REFERENCES

- [1] [BZCBZ1999] M.Z. Bleiberg, D. Zhu, K. Cole, D. Bates and W.Zhan, "Developing an integrated library decision support data warehouse", Kansas State University, Manhattan, 1999 Systems, Man, and Cybernetics, IEEE SMC '99 Conference Proceedings. 1999.
- [2] [CR2004] Cassia Blondet Baruque, Rubens Nascimento Melo, "Developing Digital Libraries Using Data Warehousing and Data Mining Techniques", PUC-RioInf. MCC28/04 July 2004.
- [3] [EA2011] Elaheh Homayounvala, Ammar Jalalimanesh, "Promoting Research Collaboration Based on Data Mining Techniques in Library Information System", International Journal of Information Technology and Business Management", Vol.8, No. 1, Dec'2012
- [4] [LGB1999] Larence S., Giles C.L., Bollacker K., "Digital Libraries and Autonomous Citation Indexing", IEEE Computer 32(6) - 1999 pp 67-71
- [5] [M2007] Mo Li, "Application of web-based data mining technology in digital libraries", journal of Academic Library and Information Science, Vol. 25, pp 44-46, August 2007
- [6] [M2011] Mei Zhang, "Application of Data mining Technology in Digital Library", Journal of Computers, Vol 6. No. 4, April 2011
- [7] [MMGP2004] Mudhol, Mahesh & Gowda, Purushothama, "Data Mining in the process of Knowledge Discovery in Digital Libraries", INFLIBNET, Ahmadabad, 2004.
- [8] [PPU2004] K Prakesh, Prem Chand and Umesh Gohel, "Applications of Data Mining in Library and Information Services", 2nd Convention PLANNER – 2004, Manipur University, Imphal 4-5 Noverber, 2004, pp 169-177
- [9] [RYKDY 2008] Ruey-Shun Chen, Yung-Shunn Tsai, K. C. Yeh, D.H. Yu and Yip Bak-Sau, "Using Data Mining to Provide Recommendation Service", WSEAS Transactions on Information Science & Applications, ISSN : 1790-0832, Issue 4, Vol 5,2008
- [10] [Z2007] Jun Zhou, "Design for Personalized Service in the Digital Library Based on Data Mining", Research in Library science. No. 3, pp 15-17, March 2007.