Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 22 (2018) 550–557

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, 5-6 October 2017, Tirgu-Mures, Romania

FUCL mining technique for book recommender system in library service

Pijitra Jomsri[a,]*

[a]*Department of Information Technology,Suan Sunandha Rajabhat University,Bangkok, 10300 ,Thailand*

Abstract

Recommender systems are important tools in library websites that assists the user to find the appropriate books. With the rapid development of internet technologies and the number of books has varied which waste of time and difficulty for finding from library searching system. This research presents a book recommendation system for university libraries to support user interests which are related in the same topic and faculty. The main motive of this research is to develop the technique which recommends the most suitable books to users according to the faculty of the user profile with book category, and book loan or FUCL technique. This is based on the combined features of association rule mining. The results show that FUCL mining technique is suitable to apply for the recommender book tool in the library and has a higher accuracy value than other technique.

Peer-review under responsibility of the scientific committee of the 11th International Conference Interdisciplinarity in Engineering.

Keywords: Recommendation System; data mining; association rule; user profile.

1. Introduction

The use of internet to search information is likely to increase. Particularly, searching books in university library. There has been a continual development of the recommender system in library systems to increase effectiveness of information searching to meet users' satisfaction at the highest level. This technique is applied to recommend

* Corresponding author. Tel.: +66-216-011-44; fax: +66-216-011-46.
E-mail address: pijitra.jo@ssru.ac.th, pijitra.jo@gmail.com

2351-9789
Peer-review under responsibility of the scientific committee of the 11th International Conference Interdisciplinarity in Engineering.
10.1016/j.promfg.2018.03.081

information which yields more relevant results to users' needs resulting in enhancing users' satisfaction when searching for books in the library.

This research is interested in developing the book recommendation system to serve individual needs from each faculty. In addition the association rule technique was applied to identify relationships between books that the users in each faculty are interested in and the availability of books in the system in accordance with book categories and book loan. This could facilitate the users when searching for books, and give better searching results. The book recommendation system not only enhanced effectiveness of library system but also helped reduce cost of keeping. Additionally, it facilitated users when searching through a wide range of books on shelves and could develop users' reading habits.

The paper is structured as follows: related works are summarized in Section 2. The proposed recommendation mechanism is presented in Section 3. Association rule model are explained in Section 4. Consequently Experimental Setting are describes in section 5 and Experimental result is explained in Section 6. Finally, Section 7 describes conclusions and future work.

2. Related Work

Recommender systems have been proposed for recommending books in literature [1-4]. Systems make use of different information filtering approaches for recommending books depending upon the context and domain in which these systems are developed such as, LIBRA [5] is a CB book recommender system that uses book information taken from the Amazon.com web pages. User profile was learned and title recommendations using ratings provided by users. It also applies text categorization to semi-structured data that is obtained through information extraction from Amazon.com web pages. Yan-ge Ma [6] proposed a recommendation algorithm which takes the interconnected characteristic of books by content similarity into account, and considers the information attributes of one book.

The K3Rec [7] book recommender is developed for K-3 readers. It examines book content, topic suitability and other features in order to recommend books to children that best fit with reading choices and readability levels of K-3 students. K3Rec is limited in assessing the quality of items if the contents lack enough information about items being recommended. In that case, K3Rec is unable to differentiate between user likes and dislikes. Other problems associated with this method include overspecialization and new user problems.

Hybrid recommenders have been developed such as combination user demographic information with CB and CF approaches in recommending books [8]. This way sparsity and cold-start problems have also been resolved which lead to the worst predictions [9]. To cope with performance and quality related issues, NOVA [10] combines different features of CF, CB, and context-aware recommender systems. In comparison with traditional approaches, NOVA shows a comparatively greater performance and quality than others. In the design of recommender systems, several data mining techniques are used to mine interesting rules useful in recommendations from large datasets. For example, a system presented in [11] uses a k-means clustering technique to make clusters of system users by using the transaction history of users and recommending personalized book lists to users within these clusters. However, due to not considering the actual book contents, the recommended books are poor representatives of the users needs. Mobasher et al. [12] presented a system for web personalization based on association rules mining. Their system identifies association rules from page views co-occurrences based on users navigational patterns. Another system is presented in [13] that makes use of library loan records and applies association rule mining techniques to recommend books in the digital library by using the association rule mining system to make inferences and derive interesting rules. Rules are made interesting such as computer students are more likely to take interest in computer and math books rather than geography hence improves recommendation quality. This approach is limited in finding interesting rules that meet the interest of users who do not visit libraries frequently and perform transactions with the system. Up to now, most of the algorithms recommend items based on the rank or score given by users, whereas in libraries' book-loan logs, no such data was available. Yan et al. [14] discussed the users' behavior from the library book-loan log, but the study mainly concentrates on the knowledge dependency of different majors. This study discusses book recommendations based on the book loan data, but their method of association rules, and the paper, only proposed a service mode without experiment on the algorithm [15].

This paper uses different views to recommend a book by using the user profile and extends the method of association to rule that is a data mining technique by using implicit and explicit data.

3. The Proposed Recommendation Mechanism

A framework for book recommender system is as follows in Fig 1. The framework mechanism was designed in three steps:

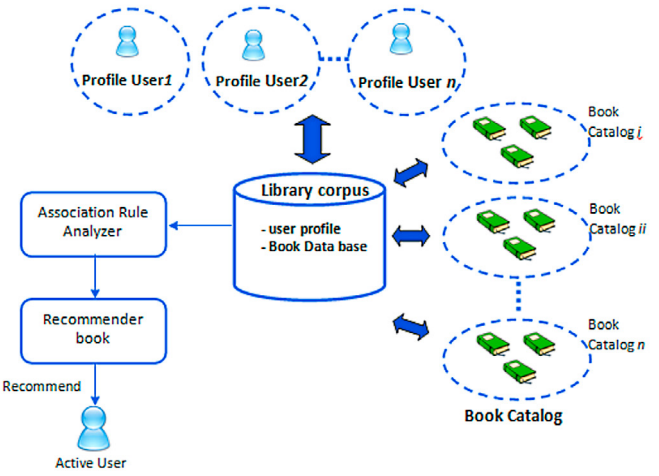

Fig. 1. Proposed Recommendation Mechanism.

3.1. Library corpus

Library corpus collects information of user bibliographic and historical data of each user loaning after processing the user loan book. The collected documents consist of book ID, book name, book category, author name, user ID, loan date, return date, and user faculty etc. In addition, the detail of each category is described as follows:

000 Miscellaneous
100 Philosophy and psychology
200 Religions
300 Social sciences
400 Languages
500 Science
600 Technologies
700 Arts & recreation
800 Literatures
900 History & geography

3.2. User Profiling

User profiling is a process of modeling a user's preferences. Assume that there are n users participating in the system, m books were loaned.

Let P be a set of all the users contained in the system; $P = \{P_1, P_2 \ldots, P_n\}$, L is a set of books loaned from library collection; $L = \{l_1, l_2 \ldots, l_m\}$, PL_{ij} is a set of user loan books by user P_i; $PL_{ij} = \{pl_{il}, pl_{i2}, \ldots, pl_{ij}\}$ and Let $E(p_i, pl_{ij})$ indicates a relationship among user P_i, with loan PL_{ij}. The user profile is defined as follows:

Definition [User Profile):

For a user p_i where $i=1, .. , n$;
Let U_i; be a user profile of user p_i.
$U_i; = \{< p_i , pl_{ij} >/ pl_{ij} \in PL \wedge p_i \in P \wedge E(p_i, pl_{ij}) = 1 \}$

3.3. Association rule Analyzer

This step is preparing and cleaning data for creating the association rule model. The relationship while users loaned book were analyzed. This paper used the association rule technique from user profiles for exploring the pattern to improve the book recommender system. Researchers explored the association of faculty, a set user, book category and book title.

The system now applies association rule mining on these transactions and finds out the books that the readers can study afterward. The support and confidence parameters are adjusted to get stronger rules. The association algorithm rules is mainly used to find out the relationships between items that occur in the database. The plan is to discover such rules that satisfy the user-specified minimum support and confidence score. Support is an indication of how frequently the items appear in the database. Confidence indicates the number of times the if/then statements have been found to be true.

Researchers expect that relation of the history of user loaning will be significant for creating the recommendation system. The detail of Association rule Mining model is as follows:

The rules tab in the form of $F,U,C,L \rightarrow T$ is applied for extracting the rules called the *FUCL Model* and $U,C,L \rightarrow T$ is applied for extracting rules called the *UCLModel*. Where F, U, C and T are disjointed item sets of user (U), Faculty (F), book category (C), book loan (L) and book title (T). For each rule of the form $F,U,C,L \rightarrow T$, when the rule fulfills the required criteria of minimum support and confidence value, it is considered useful for the purpose of recommendation. Researchers define the *supp* and *conf* as the *support* and *confidence* as follows.

$$conf(F,U,C,L \rightarrow T) = \frac{count(F,U,C,\text{L and } T)}{count(F,U,C,L)} \quad (1)$$

$$\sup(F,U,C,L \rightarrow T) = \frac{count(F,U,C,L \text{ and } T)}{count(\text{All})} \quad (2)$$

Table 1 shows examples of rules for book relation. Confidence and support values are used for rule selection. Because plenty of rules are generated, some simple concerns in rule selection include:

Select the rule with maximum confidence.
Select the rule with maximum support if confidence value is equal.
Select the rule that happens first when confidence and support values are equal.

From Table 1 can be observed the following explained rules:

- *Support of F,U,C,*L→T *is the probability that a book has in user (U), user in Faculty (F), book category (C), book loan (L) and book title (T).*
- *Confidence of F,U,C,*L→T *is probability that a book appear in book title (T) given that the user (U), user in Faculty (F), book loan (L) ,book category (C).*

Table 1. Relation Models of Book that User Loan with Confidence and Support Values.

Rule			Conf (%)	Sup (%)
science ,User 1, 000, Design and implementation of data mining tools	⇒	Next generation of data mining	73.23	1.37

3.4. Recommender Book

Recommender book process to suggest books for each user as resulting from association rule process.

4. Experimental Setting

The experimental setting is divided into two sections. Section 4.1 describes the data set, section 4.2 discusses describes evaluation metrics.

4.1. The data set

The collected documents consist of 126,521 transactions during January 2012 to February 2016. Each record in the historical loaning corpus contains: Book ID, book name, category number, return date, loan date, barcode, type of user and user id, and bibliographic information.

4.2. Evaluation Matrix

Thirty users were recruited as experiment participants. In measuring system accuracy, information retrieval classification metrics was used to evaluate the capability of the system to suggest a short list of interesting items to the user. The accuracy of precision and recall are the standard measurement for the probability that the system makes a correct or incorrect decision about the users interest. With b_x being the book recommended for user u and $D(u,b)$ is the set of recommended books, recall and are defined as Equation (3)and (4):

$$recall = (D(u,b)) = \frac{1}{|U|}\sum_{u\in U}\frac{|bp(u,b)\cap D(u,b)|}{|bp(u,b)|} \tag{3}$$

$$precision = (D(u,b)) = \frac{1}{|U|}\sum_{u\in U}\frac{|bp(u,b)\cap D(u,b)|}{|D(u,b)|} \tag{4}$$

Where

$|bp(u,b)|$ is the number of is relevant documents,

$|D(u,b)|$ is the number of retrieved documents.

$|bp(u,p)\cap D(u,p)|$ is the number of relevant documents from the number of retrieved documents.

Recall measures the percentage of interesting items suggested to the users, with respect to the total number of interesting items. Whereas, precision measures the percentage of interesting items suggested to the users, with respect to the total number of suggested items. The values precision and recall are shown in section 5. The thirty subjects that participated in the experiment were considered experts in their field. Therefore, their relevancy scores are assumed to be relatively accurate. In the study setting, each subject is assigned to investigate the book obtained from the b_x**.** The 10 books for relevancy are displayed. Finally, the subjects were asked to rate the relevancy of the book recommendation on a two-point scale: score 0 for not relevant and score 1 for relevant.

5. Experimental Results

The results of the paper were described in two sections .The first section is a result of association rule, and the second one is a result of evaluation by using precision and recall.

5.1. Result of association rule

The rule which has minimum confidence value showed that more than 60% were choose which denotes the percentage of transactions containing the user (*U*), user in Faculty (*F*), book loan (*L*) ,book category (*C*) which contain also in book title (*T*) . It is an estimation of conditioned probability. The strength rules were hold such as "Management, User 1, 000, Knowledge management" was at the 100% confidence. The rule means that user1 which loans in Miscellaneous category, book name: Knowledge management will loan in Knowledge management: theoretical concepts. Therefore, the relationship of these rules may help to recommend book to user. However, "Science, User9, 600, Need to know: social science research methods" has Confidence 37.98%. So User which loans in Technologies category, book name: Need to know: social science research methods will not loan in doing action research in your own organization. Therefore, the relationship of this rule may not help to recommend book.

Table 2. The Accuracy of Confidence Training Value and Confidence Testing Value .

Rule			***Conf (%)***	***Sup (%)***	***Rule hold***
Science,User 1, 000,Visual basic 2008	⇒	Quick Start VB. NET 2008	69.77	1.37	Yes
Science,User 1, 000,Knowledge management	⇒	Knowledge management : theoretical concepts	100	10.5	Yes
Science,User2, 600,Fundamental nutrition	⇒	Nutrition for healthy living	67.33	1.01	Yes
Science,User2, 300, Food business management	⇒	Food preparation and theory	60.00	0.86	Yes
Education,User3, 400, Education,English in scientific context	⇒	Progressive English	65.15	1.46	Yes
Education,User3, 400, TOEFL reading comprehension	⇒	TOEIC reading review-tests	66.23	0.79	Yes
Arts, User4, 700, Music : the art of listening	⇒	Music of the masters	70.71	0.78	Yes
Arts, User4, 700, Folk music and traditional performing arts of Thailand	⇒	Jazz styles : history & analysis	82.65	0.76	Yes
Humanities,User5, 100, Phychology in everyday life	⇒	General psychology	55.76	0.75	No
Humanities,User5, 100, Educational technology research	⇒	Psychology of reading	67.74	0.73	Yes
Education,User6, 800, Phra Aphai Mani II	⇒	Phra Aphai Mani III	52.38	0.72	No
Education,User6, 800, Thai usage1	⇒	Thai usage2	100	0.69	Yes
Science,User7, 000, JAVA OOP	⇒	Intro JAVA	63.00	0.67	Yes
Science,User7, 000, PHP	⇒	C++	71.21	0.68	Yes
Science,User8, 000, Digital system designs and practice	⇒	Communication skill	21.67	1.23	No
Science,User8, 600, Principles of electric circuits	⇒	Electronic devices and circuit theory	78.57	2.82	Yes
Science,User9, 300, Need to know : social science research methods	⇒	Doing action research in your own organization	37.98	1.75	No
Science,User9, 000, Java	⇒	Into JAVA	63.76	0.82	Yes
Science,User9, 000, HTML & XHTML	⇒	Flash 8	60.26	1.34	Yes
Nurse,User10, 600, Nursing process	⇒	Nursing care of the child	89.21	1.89	Yes

5.2. Result of evaluation matrix

The subject relevancy ratings for the book were recommended via the following equation (4). The experiment result is depicted in table 3. The values in each cell is the average of 30 search terms precision values. In this experiment, a comparison between *FUCLModel* with *UCLModel. FUCLModel* is the technique suggested and *UCLModel* is the comparison technique based on the user profile. The results show that the *FUCLModel* is better than the *UCLModel*. In high precision value indicates high accuracy and the result show that Precision In top-5 shows the performance better than Precision In top-1.

Table 3. shows the evolution model.

Method	Precision in top-1(%)	Precision in top-5(%)
FUCL Model	92.45%	94.52%
UCL Model	91.34%	92.00%

6. Conclusions and Future Works

This research presented a book recommendation system for the university library. The researcher analyzed user profiles which included their history of borrowing, book categories and related data. Users from same faculties borrowed books in the same category which imply that users are interested in the same book topic area. Furthermore, the findings of a model of book recommendation system developed to facilitate individual users for book searching indicated that the users satisfied with the book recommendation system. Results from the evaluation matrix suggest that the heuristic model *FUCLModel* outperforms the *UCLModel*. In further studies, the researcher will include other factors to analyze relations, apply other data mining techniques and a comparison with the other developed models.

Acknowledgements

The author would like to thank the Suan Sunandha Rajabhat University for Scholarship support. The study is not possible without the data from Academic Resource Center of Suan Sunandha Rajabhat University.

References

[1] G. Linden, B. Smith, and J. York, Amazon.com recommendations: item-to-item collaborative filtering. Internet Computing, IEEE Internet Computing (2003), pp. 76-80.
[2] X. Liu, E. Haihong, J. Song, M. Song, J. Tong, Collaborative Book Recommendation Based on Readers' Borrowing Records, International Conference on Advanced Cloud and Big Data , Nanjing, Jiangsu, China, 13-15 December, 2013, pp. 159-163.
[3] Z. Luo, Realization of Individualized Recommendation System on Books Sale, International Conference on Management of e-Commerce and e-Government (ICMeCG), Beijing, China, 20-21 October, 2012, pp.10-13.
[4] A.S. Tewari, A. Kumar, A.G. Barman, Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining. Advance Computing Conference (IACC), 2014 ,IEEE, pp. 500-503.
[5] R.J. Mooney, P.N. Bennett, L. Roy, Book recommending using text categorization with extracted information. Proc. Recommender Systems ,1998 Workshop, Technical Report WS-98-08.
[6] Y.-G. Ma, Research of Library Book Recommendation System Based on Cloud Computing, In Proceedings of the 9th International Symposium on Linear Drives for Industry Applications. 1 (2014) 549-555.
[7] M.S. Pera, Y.-K. NG., Analyzing Book-Related Features to Recommend Books for Emergent Readers. the 26th ACM Conference on Hypertext & Social Media ACM, 2015, pp.221-230.
[8] M. Chandak, , S. Girase, D. Mukhopadhyay, Introducing Hybrid Technique for Optimization of Book Recommender System. Procedia Computer Science. 45 (2015) 23-31.
[9] J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filtering recommender systems. In The adaptive web , Lecture Notes in Computer Science book series, 2007, pp. 291-324.
[10] D. Pathak, S. Matharia, C. Murthy, NOVA: Hybrid book recommendation engine. 3rd International IEEE conference on Advance Computing Conference (IACC), 2013, pp. 977-982.

[11] S. Maneewongvatana, A recommendation model for personalized book lists. In International Symposium on Communications and Information Technologies (ISCIT), Tokyo, Japan, 26-29 Oct, 2010, pp.389-394.
[12] B. Mobasher, H. Dai, T. Luo, M. Nakagawa, Effective personalization based on association rule discovery from web usage data, Web Information And Data Management, WIDM ’01, 2001.
[13] P. Jomsri, Book recommendation system for digital library based on user profiles by using association rule. Fourth International Conference on Innovative Computing Technology (INTECH), IEEE, Luton, UK, 2014, pp.130-134.
[14] F. Yan, M. Zhang, J. Tang, T. Sun, Z. Deng, L. Xiao, Users’ Book-Loan Behaviors Analysis and Knowledge Dependency Mining, WAIM 2010. LNCS. 6184 (2010) 206–217.
[15] K. Yu, A. Schwaighofer, V. Tresp, X. Xu, , H. Kriegel, Probabilistic Memory-based Collaborative Filtering, IEEE Transactions on Knowledge and Data Engineering 16(1) (2004) 56–69.