Aizu-BUS: Need-Based Book Recommendation Using Web Reviews and Web Services

Takanori Kuroiwa and Subhash Bhalla

Graduate School of Computer Science and Engineering,
University of Aizu, Aizu-Wakamatsu, Fukushima 965-8580, Japan
{m5101202,bhalla}@u-aizu.ac.jp

Abstract. Presently, there are three approaches that constitute recommender systems: collaborative filtering, content-based approach and a hybrid system. This paper proposes a complementary recommendation methodology, focusing on book recommendation. By retrieving web reviews of books using existing Web Services, an infrastructure has been developed for need-based book recommendation system. Implementation results shows that our book recommendation allows a user to eliminate irrelevant books and presents the desired books to the user from given book set. The proposed book recommender is one of the first systems in terms of focusing on meeting individuals' needs rather than calculating similarity or preferences automatically, which is adopted by the traditional recommender system.

Keywords: Book recommendation, Web review retrieval, Web Services.

1 Introduction

Recommendation systems provide users with products that they might want. However, current recommender systems such as Amazon.com's book recommender do not take each user's need into consideration. In fact, many times a user does not decide to purchase unknown recommended books promptly. This may be because in most cases, the user checks his/her own needs. For example, he/she may think "Which book is my favorite author's/publisher's?", "Which book is good for beginner?" or "If I purchase the book, are there any benefits for me?" One of the ways to resolve such questions is to check customer reviews provided by Amazon.com. Another way is to search the Web to find effective web reviews. Recently information transmissions through Blogs or virtual-community logs have been very active. Many people write book reviews on their Blogs. Therefore, if a recommender system has a facility for collecting and checking web reviews, the recommender system would be able to recommend books that meet individuals' needs.

The purpose of this study is to build an infrastructure for a need-based recommendation that supports such a facility by collecting web reviews though existing Web Service integration. Also, we presents the effectiveness of the proposed recommendation approach by providing implementation results.

S. Bhalla (Ed.): DNIS 2007, LNCS 4777, pp. 297–308, 2007.

The rest of the paper is organized as follows. The background information is described in Section 2. In Section 3, we introduce an infrastructure of our book recommendation. Section 4 discusses needs-based recommendation. In Section 5, results are discussed. Finally, Section 6 presents the summary and conclusions.

2 Background

Presently, there are three recommendation approaches that constitute recommender systems: collaborative filtering, content-based approach and their hybrid system [3]. Each approach has some drawbacks and advantages. In addition to these traditional recommendation methodologies, a new approach has been proposed. In this section, we briefly describe them.

2.1 Collaborative Filtering Approaches

The collaborative filtering approaches recommend books based on similarity between the books the user preferred in the past and the ones other users have purchased. Amazon.com's recommender system adapts this approach [7][10]. The collaborative approaches can provide books of unknown category. However, such recommender systems have to learn a user's preference from the user's previous purchases. In addition, the recommender systems require a sufficient number of other user's ratings before making a recommendation.

2.2 Content-Based Approaches

The content-based approaches recommend books based on a user's preferences [9][11][12]. The content-based approaches are appropriate for recommending books by extracting the user's preferences from the user's ratings. One of the drawbacks of such recommender systems is that they require a sufficient number of the user's ratings to make an accurate recommendation. Also, this approach may cause a loss of information problem in the course of extracting featured terms to calculate a user's interests. For example, when selecting featured terms from text, terms regarded as insignificant are eliminated. However, in terms of context, such terms may not be ignored.

2.3 Hybrid Approaches

The hybrid approaches combine the collaborative filtering and the content-based approaches [4][13]. The hybrid approaches are able to make more accurate recommendation than the above two approaches by overcoming weakness of pure approaches. However, recommended books still depends on users' own effort.

2.4 Review-Based Approach

Aciar et al. propose a new recommendation approach [1][2]. It is based on consumer product reviews written in free-form text. In this approach, each review comment is mapped into the ontology's information structure and classified into three categories: "good", "bad" and "quality" to decide whether a product is good or bad taking each commenter's skill level into consideration. The drawback is that this approach cannot be applied to all review comments. For example, some long, complicated sentences cannot be classified into any category.

2.5 Human-to-Human Approach

We have been proposed a human-to-human recommendation approach [8]. In this approach, users share a Knowledge Base (KB) and each user recommends his/her chosen books to other users like real-world borrower or lender through the web-based system. The purpose of this approach is to enhance utilization of existing books among users. In this approach, to visualize individuals' preference, featured keywords are extracted from the KB. However, to make an accurate featured keyword extraction, a sufficient number of user's ratings are required.

2.6 The Proposal

Both the collaborative filtering approach and the content-based approach have a common drawback: these require a sufficient number of users' efforts before making a recommendation. Similarly, the proposed review-based approach cannot work well in a few cases. The human-to-human approach also have some drawbacks. This may be because these approaches tend to focus on automatically processing something (such as extracting users' interests, calculating the similarities between users, analyzing review comments) rather than reflecting users' direct needs. Thus, making a recommendation based on these approaches may cause a gap between the recommended books and the ones users really want. This is because the approaches ignore users' current interests and needs. For this reason, we propose a recommender system that supports a facility for checking if recommended books meet the users' needs or not.

3 Aizu Book Utilization System (Aizu-BUS)

In this Section, we introduce an infrastructure for the proposed book recommendation system called Aizu Book Utilization System (Aizu-BUS).

3.1 System Architecture

Figure 1 depicts an architecture of Aizu-BUS. The Aizu-BUS is composed of Web User Interface (WUI), Web Service handler, XML DB handler, external resources and a Knowledge Base (KB). The WUI receives users' requests such as gathering and searching operation for books and generates query to the Web

Services. The Web Service handler searches the external resources according to the request and give the results as store/query operation. The XML DB handler stores XML data or retrieves XML data from the XML database called Xpriori [16]. The KB is used for book utilization activities: book search, book recommendation and book sharing among users. In this study, we focus on book recommendation activity.

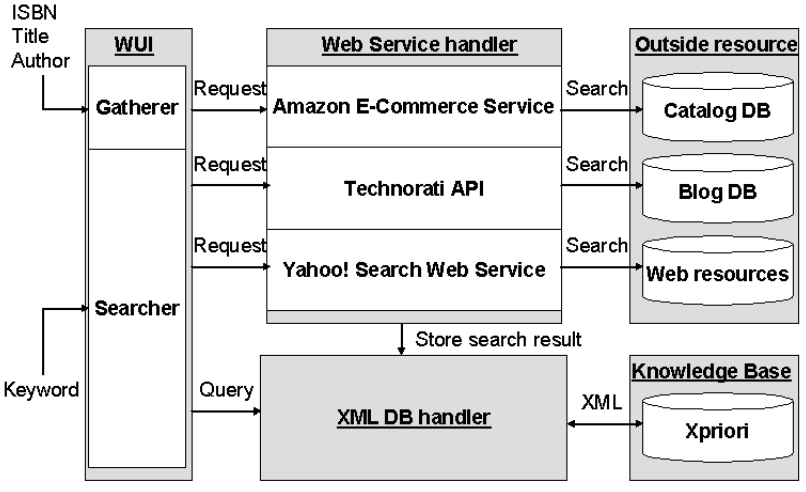

Fig. 1. System Architecture

3.2 Building Knowledge Base

Initially, data about books (available in store) is entered. When a user enters an ISBN (as shown in Figure 1), the system generates a request to Amazon E-Commerce Service [14] and searches book catalog to retrieve basic book information. The basic book information includes: image URL; ISBN; title; author name; product page URL; price; publication date; publisher name; and the number of pages and Amazon.com's sales rank. The basic book information is stored in an XML database called Xpriori. The XML database serves as a KB for the proposed book recommendation system. Technorati API [15] and Yahoo! Search Web Service [17] are applied for Web and Blog information retrieval.

3.3 Book Rating

Aizu-BUS provides a book rating method considering the fact that different users may have different rating scale [5][6]. The book rating approach given by the following formula depends not on the score of each book but on the context of each book review.

$$r = \frac{N_p - N_n}{N_p + N_n} \times (r_{ave} - 1) + r_{ave} \tag{1}$$

Table 1. Book Rating

r	Evaluation
1	Poor
$1 < r \leq 2$	Bad
$2 < r \leq 3$	So-so
$3 < r \leq 4$	Nice
$4 < r \leq 5$	Excellent

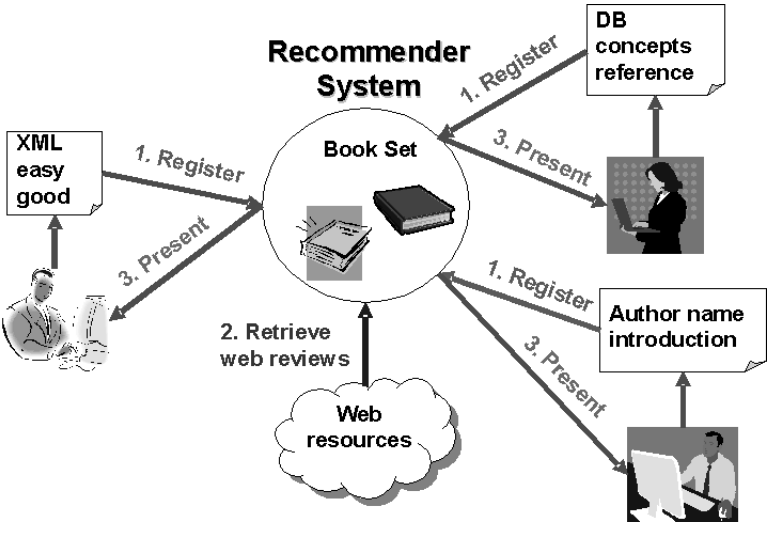

Fig. 2. Need-Based Book Recommendation

where r_{ave} is the average rating, N_p and N_n represents the positive review count and the negative review count respectively. The value of r ranging from 1 to 5 represents book ratings. The larger the value of r is, the better evaluation is. Table 1 shows the definition of book rating in this study. For example, if the number of the positive reviews of a book (for example, review comments are saying "good" or "easy") equals to 3 and the number of the negative reviews of the book (for example review comments is saying "bad" or "difficult") equals to 1, r will be 4, which represents the book is "Nice". Figure 4 and 5 depict the proposed book rating. Suppose that each book review is collected manually. In case two positive reviews (that contain "excellent" and "useful" in the review comments) are added to Knowledge Base, the book rating (r) become excellent (5) as shown in Figure 4. However, if one negative review (that contains "worst") is added, the book review (r) is reduced to "nice" (3.66) as shown in Figure 5.

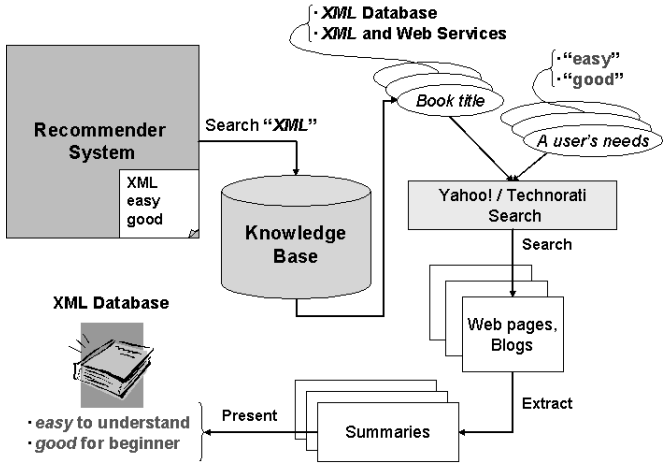

Fig. 3. Web Review Retrieval

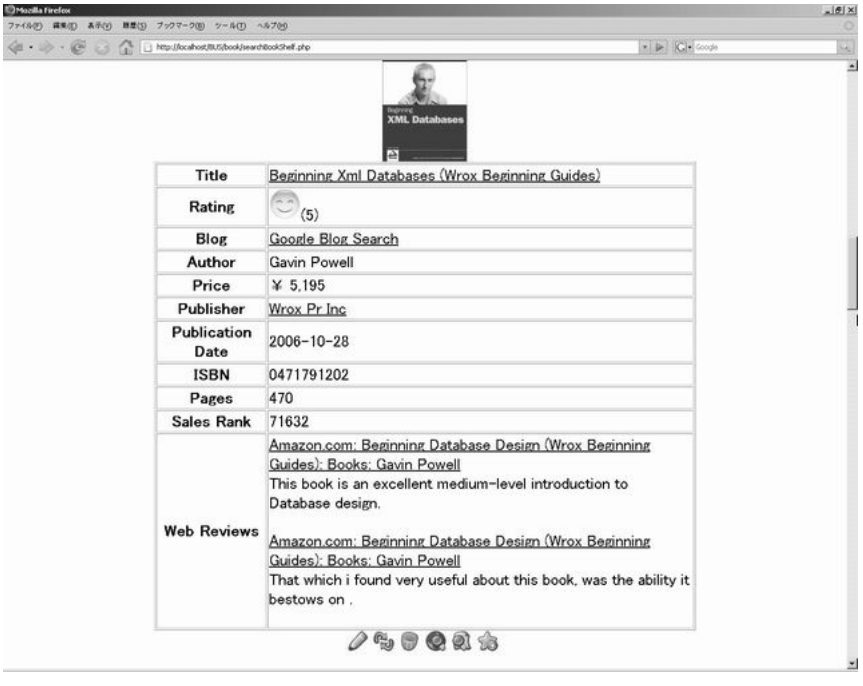

Fig. 4. Review-based Book Rating [Excellent]

4 Need-Based Book Recommendation

In this Section, we propose a need-based recommendation as a facility of the Aizu-BUS. The need-based recommendation approach corresponds users' direct needs

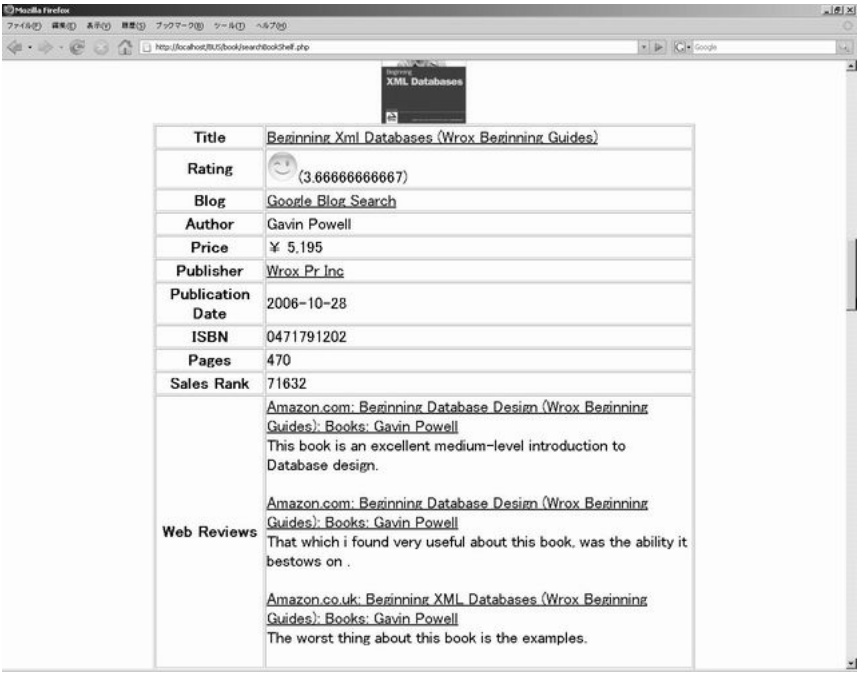

Fig. 5. Review-based Book Rating [Nice]

with book reviews available on the Web using the existing Web Services. As shown in Figure 2, the need-based book recommendation follows Step 1 to 3 below.

1. Needs registration;
2. Web review retrieval; and
3. Book recommendation

4.1 Needs Registration

First, each user registers his/her needs into the recommender system. The format of contents is as follows 1) a representative keyword of his/her interests (such as an interested category name, a favorite author/publisher name), 2) his/her first need (such as "easy", "concepts" or "introduction"), and 3) his/her second need (such as "good" or "reference"). These registered needs are used as keywords for narrowing down the search result when searching the Web/Blogs using Yahoo!/Technorati Web Services.

4.2 Web Review Retrieval

Once the needs are registered, the recommender system automatically assembles queries and searches the Web/Blogs using Yahoo!/Technorati Web Services.

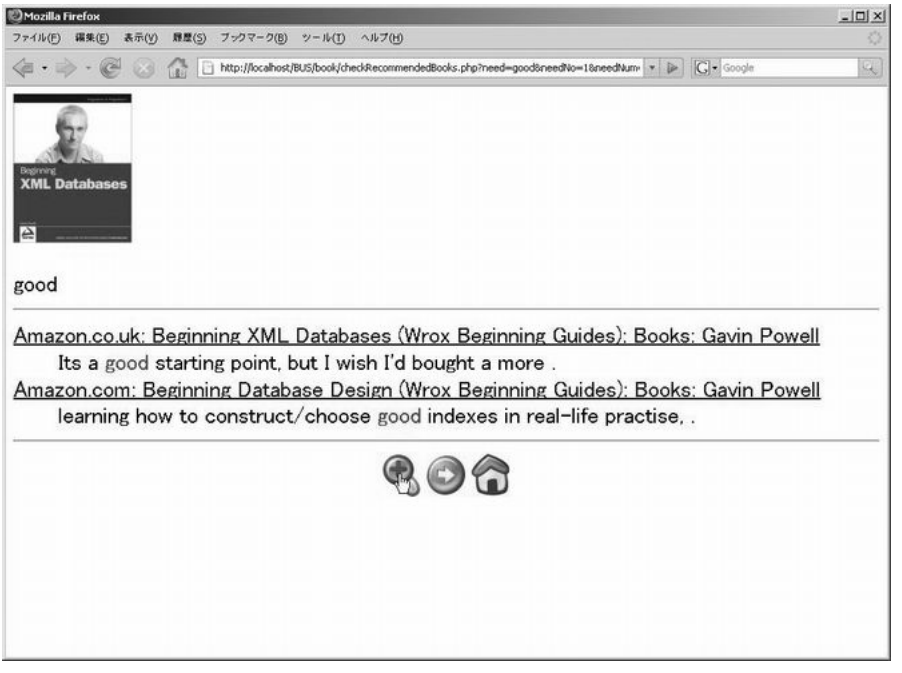

Fig. 6. Recommended book 1 containing "good"

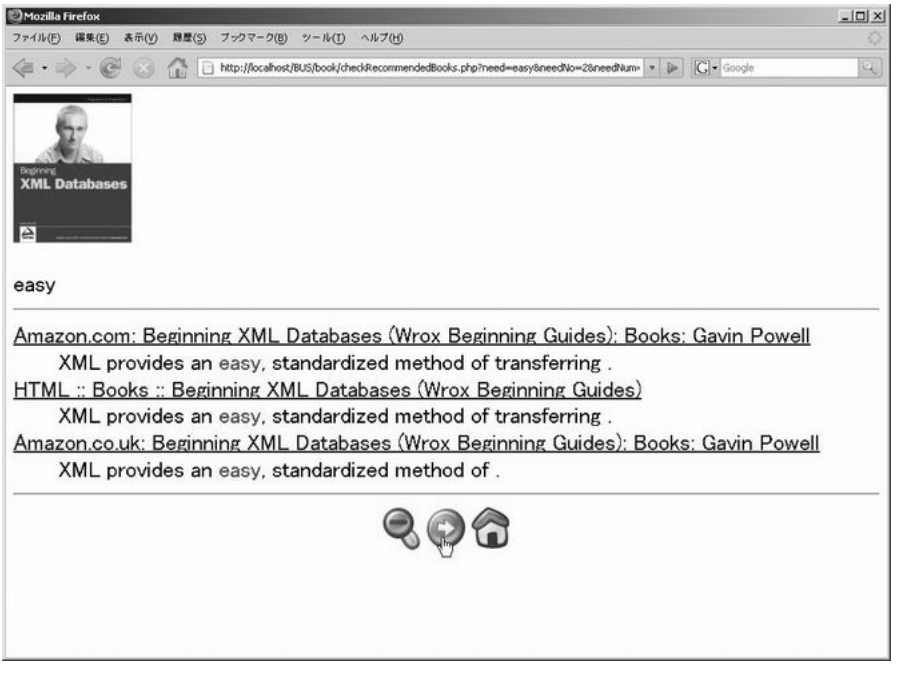

Fig. 7. Recommended book 1 containing "easy"

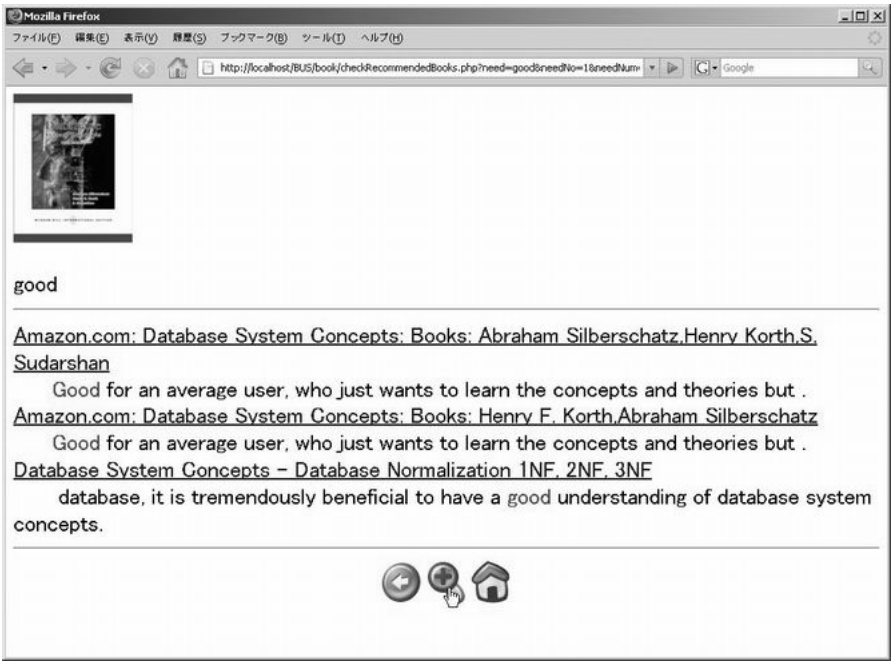

Fig. 8. Recommended book 2 containing "good"

Figure 3 shows the procedures of Web review retrieval. First, recommender system searches the Knowledge Base by using a user's first need (that is, for example, a representative keyword such as "XML") and book titles containing the keyword are retrieved. Then, the system automatically assembles the query to search web reviews by combining a selected book title and the user's next need "easy". By using Yahoo!/Technorati Web Service, short summaries containing the assembled query are extracted from retrieved Web pages/Blogs. Finally, the extracted summaries are presented to the user.

4.3 Book Recommendation

After retrieving the web reviews according to the above method, our book recommender system presents recommended books to a user one by one according to Amazon.com's sales rank. Currently, the web review retrieval and recommended book presentation are performed in real-time. In other words, a user can trigger those two activities. When a user requests to browse recommended books from given book set (such as other user's library) , the recommender system searches the web reviews using Web Services and presents the selected books (that meet his/her needs) to him/herself. The implementation details and results are discussed in the next section.

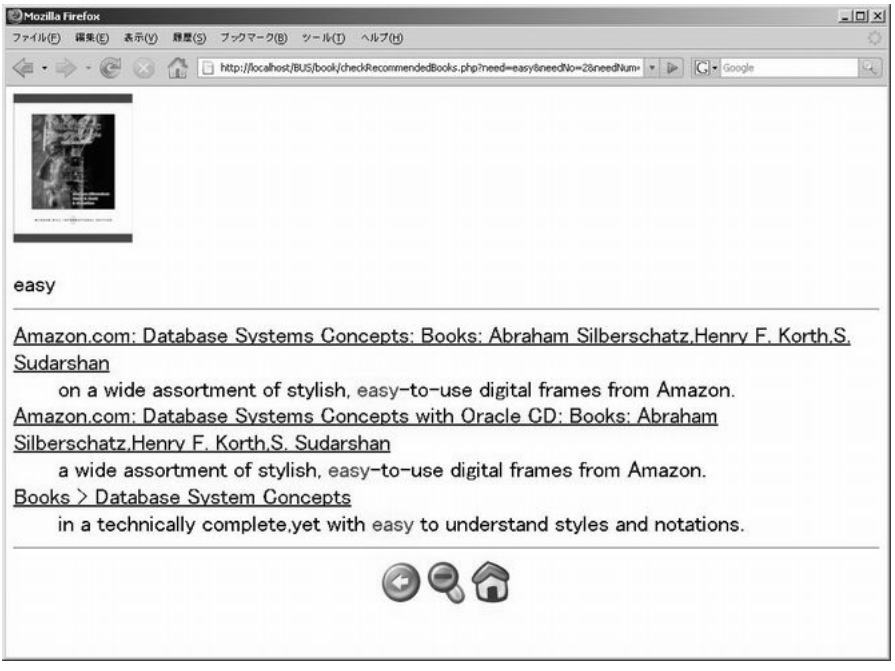

Fig. 9. Recommended book 2 containing "easy"

5 Implementation Details

A prototype system was implemented on Windows 2000 using Apache 2.0.5 and PHP 5.1.4. To test our book recommendation facility, we collected 100 books from Amazon.com. Assume that a user has registered his/her needs like "Database", "good" and "easy". Also, note that the top three pages of Yahoo!/Technorati search are retrieved as web reviews of a selected book.

Figure 6 to 9 show the recommended books and the review comments retrieved by Yahoo! Web Service. Most reviews are from Amazon.com related sites. This is because Blog search results have not been taken into account. Technorati API (that is used for retrieving book reviews from Blogs) is unsuitable for retrieving web contents written in space delimited language like English because it does not support a perfect matching search (at this time). However, it is appropriate for retrieving Japanese contents. Therefore, we apply Technorati API to Japanese book review retrievals.

6 Summary and Conclusions

This paper proposed a need-based book recommendation using Web reviews. Implementation results show that existing Web Service integration contributes

to retrieving book reviews available on the Web and corresponding the web reviews with each user's needs. This recommendation method is not an alternative approach to traditional approaches such as the collaborative filtering and the content-based approaches, but a complementary approach to such approaches. We believe the proposed book recommender is one of the novel systems in terms of taking users' needs directly into consideration and focusing on meeting the needs rather than automatically predicating recommended books, as in the case of the traditional recommender systems.

For future work, we will investigate the following things: coordination with traditional recommendation approaches and web review clustering to automatically classify web reviews into the positive ones and the negative ones.

References

1. Aciar, S., Zhang, D., Simoff, S., Debenham, J.: Recommender System Based on Consumer Product Reviews. In: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 719–723 (2006)
2. Aciar, S., Zhang, D., Simoff, S., Debenham, J.: Informed recommender: basing recommendations on consumer product reviews. IEEE Intelligent Systems 22(3), 39–47 (2007)
3. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17(6), 734–749 (2005)
4. Balabanovic, M., Shoham, Y.: Fab: Content-Based, Collaborative Recommendation. Communications of ACM 40(3), 66–72 (1997)
5. Jin, R., Si, L., Zhai, C.: Preference-based Graphic Models for Collaborative Filtering. In: Proceedings of Nineteenth Conference on Uncertainty in Artificial Intelligence (August 2003)
6. Jin, R., Si, L., Zhai, C., Callan, J.: Collaborative Filtering with Decoupled Models for Preferences and Ratings. In: Proceedings of the Twelfth International Conference on Information and Knowledge Management, pp. 309–316 (November 2003)
7. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., Riedl, J.: GroupLens: Applying Collaborative Filtering to Usenet News. Communications of the ACM 40(3), 77–87 (1997)
8. Kuroiwa, T., Bhalla, S.: Dynamic Personalization for Book Recommendation System using Web Services and Virtual Library Enhancements. In: CIT 2007. IEEE 7th International Conference on Computer and Information Technology, University of Aizu, Japan (2007)
9. Lang, K.: NewsWeeder: Learning to Filter Netnews. In: Proceedings of the 12th International Conference on Machine Learning, pp. 331–339 (1995)
10. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Computing, 76–80 (January/February 2003)
11. Mooney, R.J., Roy, L.: Content-based Book Recommending Using Learning for Text Categorization. In: Proceedings of the 5th ACM conference on Digital Libraries, pp. 195–204. ACM Press, New York (2000)
12. Pazzani, M., Billsus, D.: Learning and Revising User Profiles: The Identification of Interesting Web Sites. Machine Learning 27, 313–331 (1997)
13. Pazzani, M.: A Framework for Collaborative, Content-Based and Demographic Filtering. Artificial Intelligence Review , 393–408 (December 1999)

14. Amazon Web Service. http://aws.amazon.com/
15. Technorati API. http://technorati.com/developers/api/
16. Xpriori (XMS), http://www.xpriori.com/
17. Yahoo! Search Web Service, http://developer.yahoo.com/search/