IIAI Open Conference Publication Series
IIAI Letters on Informatics and Interdisciplinary Research
Vol. 003, LIIR054
DOI: https://doi.org/10.52731/liir.v003.054

A Case Study on the Effectiveness of Recommendation Algorithms for Sake Review Website

Tessai Hayama *, Daichi Niiyama *

Abstract

Recommendation systems have been used in various online businesses and have been recognized as useful for promoting product purchase and information browsing. However, the system design policies for each target service remain insufficiently clarified. One reason for this is that few case studies in practical environments exist since only the provider of the commercial platform obtains and confirms operational data of the recommendation algorithm and the resulting user behavior. Therefore, in this study, we investigate how some recommendation algorithms affect user behavior in online services for the Japanese Sake system. In the investigation, one of the four recommendation algorithms was assigned to each of the 858 users who logged into the system during 80 days. For each recommendation algorithm, the number of system logins, transitions to the sake description pages, sake brands registered in the sake to-drink list, and posted review comments to each sake brand were analyzed from the system history. Our results show that the recommendation system with association analysis was effective in recommending sake brands for its review sites, considering individual preferences and general popularity.

Keywords: Case study, Collaborative Filtering Recommender Systems, Experimentation.

1 Introduction

The number of social networking services (SNS) users who share their daily habits and evaluations of purchased products on SNS, such as Twitter and Facebook is increasing [1,2]. The shared comments, reviews, and behavioral data of users on the services include their preferences and tendencies [3]. Therefore, recommendation systems using a large amount of shared data on SNSs can be used to support product search activities of users and increase their purchasing motivation [1]. Notably, commercial recommendation systems for various products and services have been actively adopted [4,5,6]. However, the design policy of product and service recommendation systems has not yet been identified. One reason for this is that, since only the commercial platform provider controls the system functions and collects the user behavior data for the applied recommendation systems, only a few case studies analyzed the data and confirmed the effectiveness of the recommendation system in practical environments.

Previous studies on recommendation systems show that some researchers have analyzed some shared data from commercial platforms for wine and online games. The analysis of the SNS data from 9.2 million wine brands and 29.9 million reviews showed the tendency of user preference of wine in terms of the origin and age of the wine [7,8]. Although the tendency of user preference is useful for designing user models in practical recommendation systems, user behavior affecting

* Nagaoka University of Technology, Niigata, Japan

the commercial effectiveness of the recommendation system remains unclarified. On recommendation systems for online games, they investigated how the recommendation algorithm of the system was affected by user behavior by comparing some recommendation algorithms for about 22,000 users [9]. The results showed that the recommendation algorithm affected the increase in the number of item views, downloads, and sales. Furthermore, the recommendation algorithm with item-based collaborative filtering was effective for all evaluation items. The results are significant for designing commercial recommendation systems. However, more practical studies on the effectiveness of recommendation algorithms for different targets are needed to generalize the findings.

In this study, we examined the influence of recommendation algorithms on users in a sake review site. To achieve this, we analyzed the effect of four recommendation algorithms on user behavior using the system history of Sakeai, an online sake review service with 858 users and 23,000 registered items, for a month. The results showed that the recommendation algorithm of relevance analysis was useful in recommending unknown and interesting sake brands while considering personal preferences and general popularity as the target of recommendation.

2 Previous Studies

This research seeks to understand the effects of sake recommendation algorithms on user behavior. In this section, we will introduce studies on the comparison and evaluation of various recommendation algorithms based on customer datasets, and discuss the recommendation algorithms to be investigated in this study.

Huang et al. [10] compared several recommendation algorithms on online clothing retailers and book datasets, and the movie rating dataset from the MovieLens project. The results showed that the recommendation system with item-based collaborative filtering and link structure analysis had the highest accuracy on the movie rating dataset and clothing retail and book datasets, respectively. The recommendation system with the matrix factorization algorithm showed accuracy on average for all datasets, whereas the one with the user-based collaborative filtering showed lower accuracy than that of the other recommendation algorithms. The recommendation system with the Top-N algorithm gave high accuracy results for each dataset and recommended different items than the one with collaborative filtering algorithms.

Pradel et al. [11] compared several collaborative filtering recommendation algorithms using a dataset of the purchase history of a home improvement store and obtained that the recommendation system with association analysis had the highest accuracy. For complete purchase histories of users, the recommendation system with matrix factorization realized improved accuracy, stating frequent feedback from users as an important factor for improvement.

Benouaret et al. [12] conducted a comparison experiment of several recommendation algorithms using a dataset of customer purchase history provided by TOTAL, a French oil company. The results showed that the recommendation system with item-based collaborative filtering obtained the best accuracy with the evaluation and ranking indices, followed by that with matrix factorization and association analysis in the order of accuracy. Additionally, results showed that the accuracy of the recommendation system depends on the time of acquiring the training dataset. For example, it was reported that it was effective to use the training datasets within the expiration date for the recommendation system with item-based collaborative filtering, the most recent dataset for that with association analysis, and about 6 months old dataset for that with matrix factorization. Furthermore, all recommendation algorithms gave better accuracy than simply recommending the most popular items, such as the Top-N algorithm. Jannach et al. [9] conducted a comparative study of recommendation systems with one of several recommendation algorithms using the number of downloads, sales, and page transitions in a game recommendation system for mobile internet applications. The results showed that the recommendation system with content-based filtering obtained the highest scores for the number of items registered in the recommendation list and downloads, and the recommendation systems with the other algorithms also outperformed the Top-N algorithm. Additionally, the recommendation system with content-based filtering showed the best results in improving the overall sales rate because the recommendation target of games was a highly polarized preference of users. Moreover, this research can investigate the recommendation system in a practical environment since investigating user behavior toward the recommendation result is easier to fit into the system design guidelines for commercial use. To generalize the design guidelines for commercial recommendation systems, conducting a practical investigation of recommendation algorithms for different targets is necessary. However, since such investigations in a practical environment can be conducted only by the commercial platform providers, few case studies have been conducted.

Previous studies investigating the influence of recommendation algorithms on various datasets showed that the dominant recommendation algorithms depend on the recommendation target and dataset. Furthermore, it is possible to investigate the user's behavior toward recommendation algorithms in a real-world environment, thereby greatly contributing to the design of commercial recommendation systems. In this study, we compare and investigate several recommendation algorithms using commercially operated online sake review services to identify the effectiveness of recommendation algorithms in a practice environment.

3 Approach to Investigate

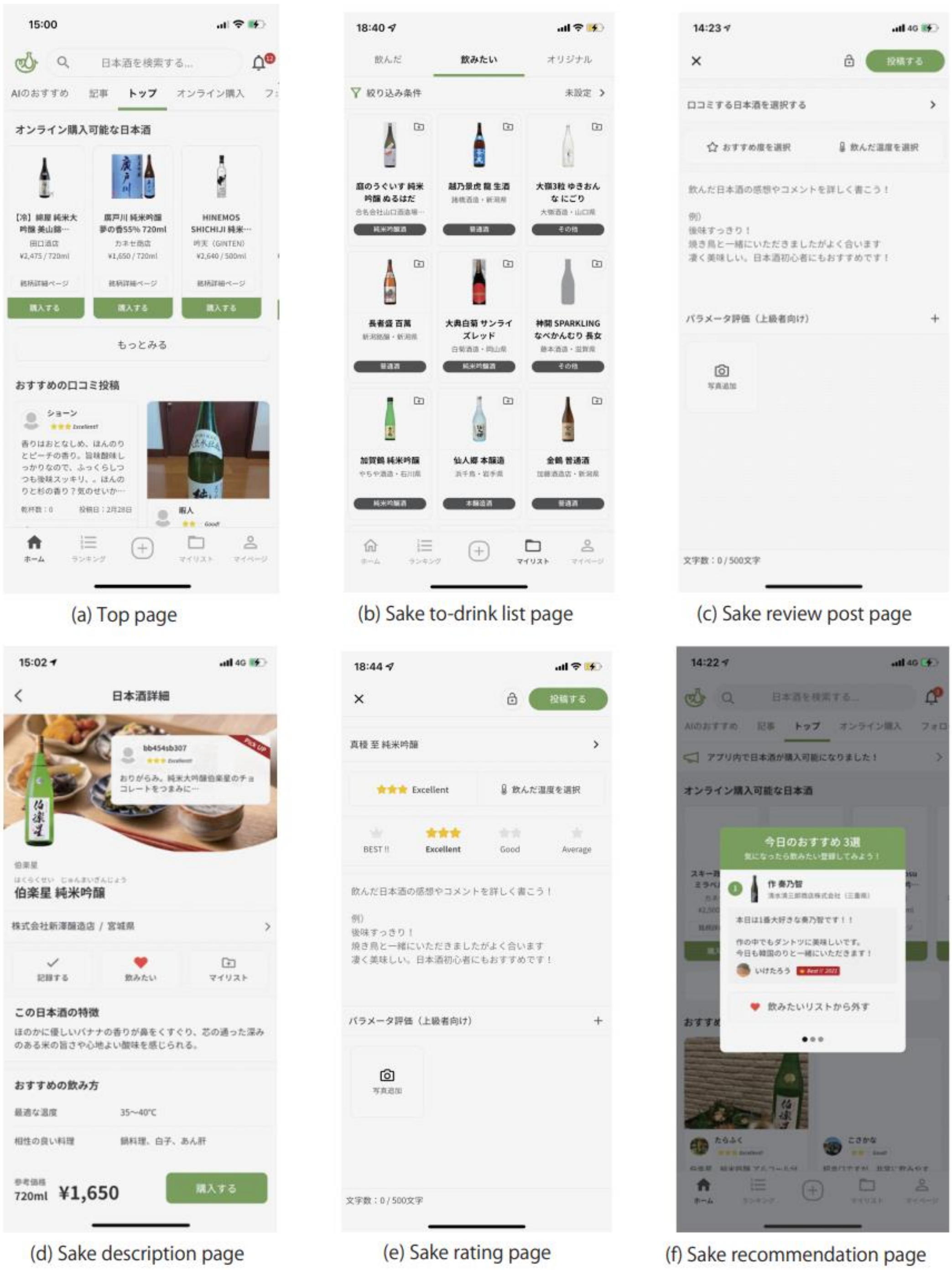
3.1 Target Environment

In this study, we use the sake review site "Sakeai[†]", operated by one of the authors since May 2020, to analyze how user behavior was influenced by a recommendation algorithm of sake brands. "Sakeai" is a smartphone application for sharing sake ratings and reviews, and as of December 2021, it has 23,800 annual accesses, 9,300 registered users (of which 800-1000 are active users accessing the application at least once a month), 24,648 registered sake brands, and 18,018 reviews. It also includes a database of composition and description published by sake breweries for each sake brand; 'Pure rice sake or brewing alcohol-added sake','Daiginjo, Ginjo, and others', 'Type of rice used (Koji rice) or not', 'Type of rice used (Kake rice) or not', 'Rice polishing ratio (Koji)', 'Rice polishing ratio (Kake)', 'Types of Yeast', 'Sake content(%)', 'Alcohol content(%)', 'Degree of acidity', and 'Degree of amino acidity'.

A user can launch "Sakeai" installed on a smartphone from the top screen. The reviews posted to the system and the sake brands registered in the user's sake-to-drink list are displayed. The transited page from the web link assigned to the sake brand list is described in the details of the sake brand, as shown in Fig. 1 (a). Furthermore, there is a function to rate each sake brand, as shown in Fig. 1 (b). From Fig. 1 (c), the system recommends three sake brands upon launch,

[†] https://sakeai.co.jp/

using a recommendation algorithm based on the user's sake-to-drink list and ratings to their sake brands. The system history of "Sakeai" mainly includes the user's login date and time, pages browsed in the system, access dates and times of the pages, and rating and review data for each sake brand. The system manages the sake-to-drink list created by each user, with reviews and ratings that the user has given to each sake brand.



(a) Top page (b) Sake to-drink list page (c) Sake review post page

(d) Sake description page (e) Sake rating page (f) Sake recommendation page

Figure 1: Screenshots of the interface for the sake review service "Sakeai".

3.2 Evaluation Method

We will investigate user behavior using system history by applying and operating some recommendation algorithms on the target platform described in Section 3.1. The recommendation algorithms adopted in the investigation are “content-based filtering,” “item-based collaborative filtering,” “association analysis,” and “matrix factorization,” similar to the representative algorithms from previous studies described in Section 2. To evaluate the user behavior influenced by each recommendation algorithm, the number of sake brands registered in the sake to-drink list, page transitions to the description page of each sake brand, and review posts for each sake brand are used, indicating that users operate the system by viewing the recommendation results during the investigation. The number of sake brands registered in the sake to-drink list, transitions to the sake description page, and sake review postings for a sake brand provides indicators of the user’s interest, unknown and possible interest, and the user’s known interest, respectively. A recommendation system that provides high value services should recommend unknown items of interest to each user.

The implementation method of each recommendation algorithm used in this investigation is shown in (1)-(4) below.

(1) Content-based Filtering

The recommendation system with a content-based algorithm recommends items similar to those the user already showed interest in [13]. For the sake brand recommendation system in the investigation, the similarity between the sake brands already registered/rated by a user and the other is calculated using Equation (1) and sake brand database. The sake brands are recommended in order of similarity.

$$sim(s, f) = \sum_{i=1}^{n} (s_i - f_i)^2. \tag{1}$$

Here, $sim(s, f)$, s_i, and f_i represent the similarity between a sake brand s registered/rated by a user and a sake brand f, the i-th component item variable of sake brands s and f, respectively. For each component variable, the normalized values between zero and one were used for components of rice polishing ratio, sake content, degree of acidity, alcohol content, and amino acid content, while a 0/1 dummy variable was used for the other components. Here n is the number of components of the sake brand, which is 11 in the database. Therefore, the closer the $sim(s,f)$ is to zero, the higher the similarity of the sake brands. Note that the system recommends sake brands in the order of decreasing similarity.

(2) Item-based Collaborative Filtering

The recommendation system with item-based collaborative filtering estimates the ratings of unrated items using users' past ratings of items and the similarity between items each user rated similarly. From the results, items with the highest predicted ratings for each user are recommended in that order [14]. For the sake brand recommendation system, the similarity between sake brands i and j was calculated using

Equation (2), and the brand with the highest similarity to the sake brands previously rated by a user in the system was recommended.

$$sim(i,j) = \frac{R_i^T R_j}{\sqrt{R_i^T R_i}\sqrt{R_j^T R_j}}. \tag{2}$$

Here *sim(i,j)* represents the similarity between sake brands i and j, R_i is the i-th column of the matrix R of sake brand ratings, and R^T is the transpose matrix of matrix R.

(3) Association Analysis

The recommendation system with association analysis extracts valuable pairs of items among all items and recommends users' unrated items highly related to items previously rated [15]. For the sake recommendation system, the ratings of the user's unrated brands were predicted for each user using Equation (3) and the history of the user's ratings of sake brands. Thus, sake brands with the highest predicted ratings were recommended to each user in that order.

$$score(u,j) = \sum_{i \in H_u} conf(i \Rightarrow j). \tag{3}$$

Here $score(u,j)$, H_u, and $conf(i \Rightarrow j)$ represent the predicted rating of unrated sake brand j of user u, list of sake brands rated by user u, and confidence score of liking sake brand j and i, respectively. The confidence score is calculated using Equation (4).

$$conf(i \Rightarrow j) = \frac{R_i^T R_j}{\sum_{j=1}^{n} |R_j|}. \tag{4}$$

Here, R_i represents the evaluation matrix in column i, containing the evaluation values of all users in the sake brand i.

(4) Matrix Factorization

The recommendation system with matrix factorization calculates the predicted score of unrated items of each user by decomposing it into a product of matrices representing the users and rating of the items, respectively. It recommends high-rated items among the unrated items for each user [12].

For the sake brand recommendation system, the matrices representing the users and evaluations of each sake brand are multiplied using Equation (5), and the sake brands are recommended in the order of the highest rating.

$$\hat{p}_{ui} = x_u^T y_i. \tag{5}$$

Here, $\hat{p}_{ui}$, $x_u{}^T$, and y_i represent the matrix of the estimated rating of the sake brand i for user u, transpose matrix component of user u, and matrix component of the rating of the sake brand *i*, respectively. The sake brand ratings were regularized using the alternating least-squares method to suppress outliers in each rating.

4 Evaluation Results

In the investigation, we collected data while operating "Sakeai," an online sake review service, from November 11, 2021, to December 30, 2021. During this period, each user was randomly assigned to one of the following recommendation algorithms: content-based filtering, item-based collaborative filtering, association analysis, and matrix factorization, and was recommended some sake brands with their images on the system interface. The number of users and user system access for each recommendation algorithm during the period are shown in Table 1 and Fig. 2, respectively.

Table 1: Number of Users for Each Recommendation Algorithm Assigned in the Investigation

Algorithm	Number of Users
Contents-based Filtering	190
Association Analysis	200
Item-based Collaborative Filtering	228
Matrix factorization	240

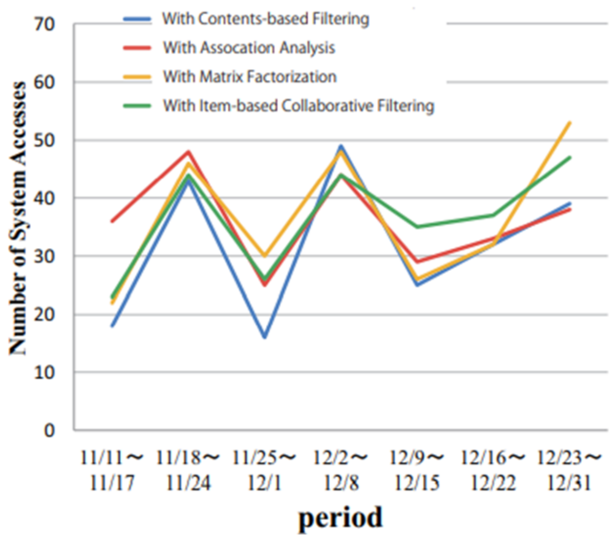


Figure 2: Number of User System Access for Each Recommendation Algorithm Assigned in the Investigation.

From Table 1, 858 target users were assigned to one of the recommendation algorithms: 190 for content-based filtering, 200 for association analysis, 228 for item-based collaborative filtering, and 240 for matrix factorization. The difference in the number of users assigned to each recommendation algorithm follows from the exclusion of users who did not log in during the target period. From Fig. 2, the users assigned to each recommendation algorithm had a similar trend in the number of accesses during the investigation.

Next, the total sake brands registered in the sake-to-drink list, page transitions to each sake brand description, and sake review postings for each recommendation algorithm are shown in Figs. 3-5, respectively. Additionally, the numbers of the sake brands registered in the sake-to-drink lists, page transitions to each sake brand description, and review postings for each sake brand in the popularity rankings for each recommendation algorithm are shown in Figs. 6-8, respectively. About 10% of the users who logged into the system during the period used the sake-to-drink list function. The popularity ranking of the sake brands was created using users' ratings of sake brands in the

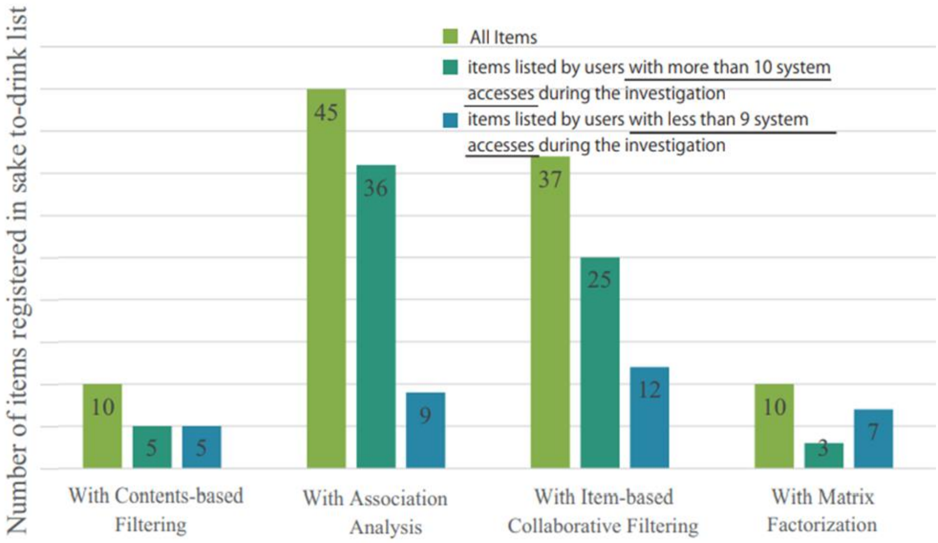


Figure 3: Total Sake Brands Registered in the Sake To-drink List for Each Recommendation Algorithm.

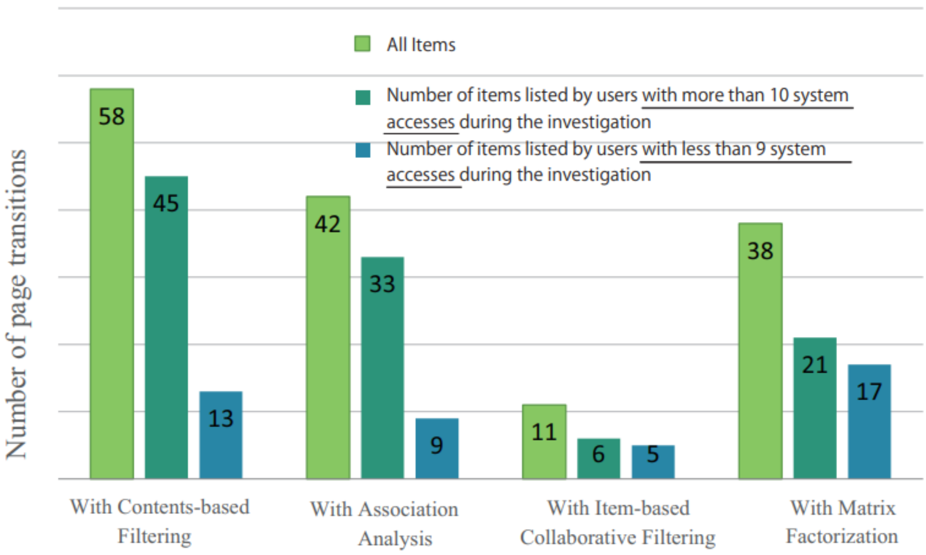


Figure 4: Total Transitions to Sake Brand Description Page for Each Recommendation Algorithm.

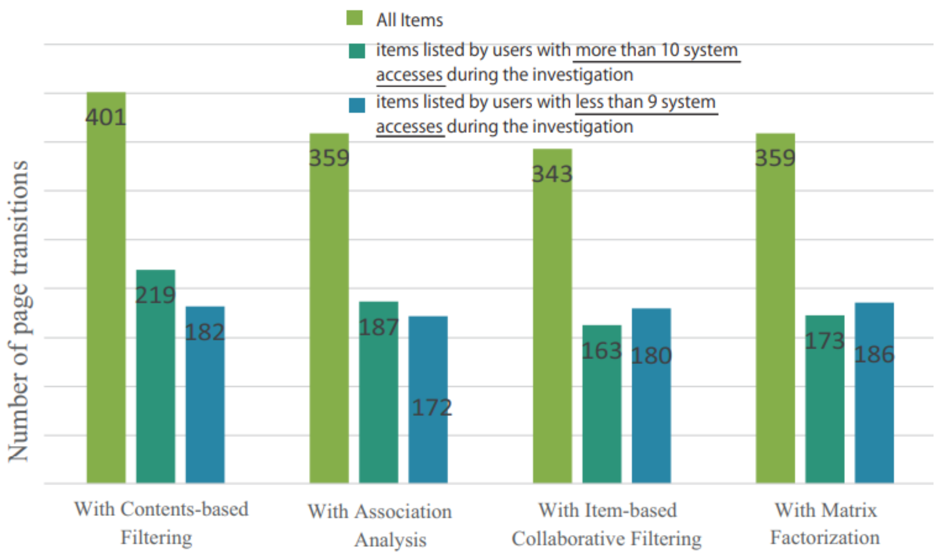


Figure 5: Total Review Posts to Each Sake Brand for Each Recommendation Algorithm.

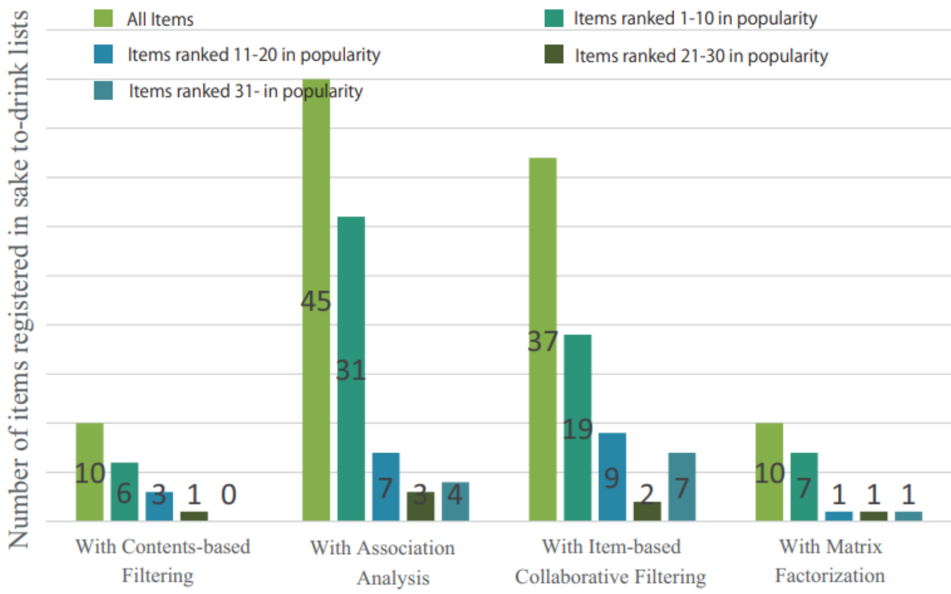


Figure 6: Number of Sake Brands Registered in the Sake To-drink List for Each Recommendation Algorithm in the Popularity Ranking.

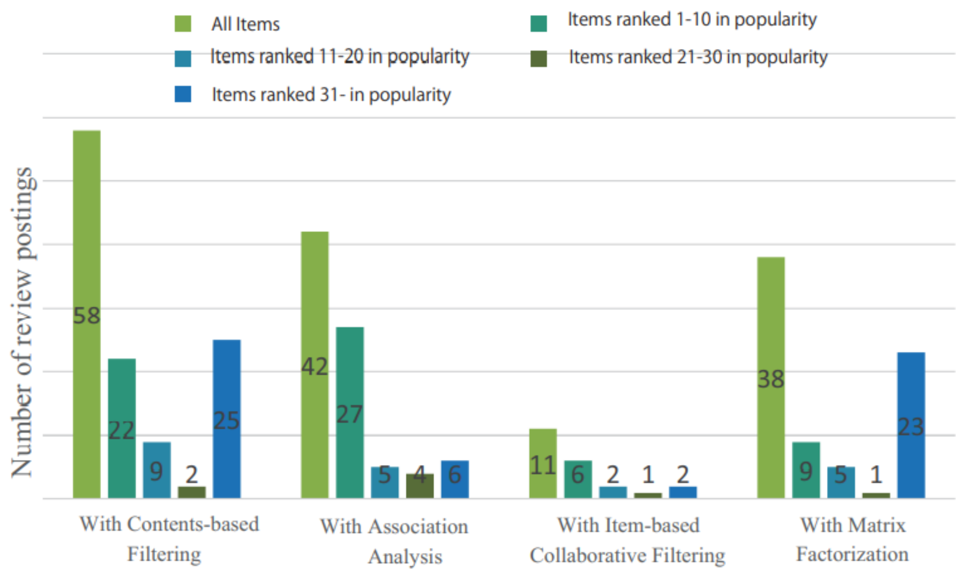


Figure 7: Number of Transitions to Sake Brand Description Page for Each Recommendation Algorithm in the Popularity Ranking.

Figure 8: Number of Review Posts to Each Sake Brand for Each Recommendation Algorithm in the Popularity Ranking.

For the sake brands registered in the sake-to-drink list, the recommendation system using the association analysis and item-based collaborative filtering registered 45 and 37 sake brands, respectively, which were more than the 10 recorded for each content-based filtering and matrix factorizations. Additionally, the recommendation system using item-based collaborative filtering showed that users who activated the system more than 10 times during the period accounted for 0.67 of the total sake brands registered in the sake-to-drink list.

For the sake description page transitions, the recommendation system using content-based filtering, association analysis, and matrix factorization resulted in higher transitions (58, 42, and 38, respectively) than the one with item-based collaborative filtering (11). Additionally, users who activated the system more than 10 times during the period accounted for 0.78 and 0.79 of the total page transitions in systems with content-based filtering, association analysis, and matrix factorization, respectively. However, users who activated the system with item-based collaborative filtering and matrix factorization more than 10 times during the study period were the same or the small difference in the page transitions from the users who activated the system less than 9 times. It showed that no effect was recorded on the number of times a user activated the system.

For review postings on sake, the recommendation system with content-based filtering obtained 401, which was 42 more than that of other algorithms. Each user posted 2.11 posts on average during the 80 days. The recommendation system with the other three algorithms obtained the sake review postings of 343 and 359, and there was little difference in the number of review postings between the three algorithms, even after considering the number of each user assigned to each algorithm. For all algorithms, the ratio of users who activated the system more than 10 times those who activated the system less than 9 times during the period ranged from 0.45-0.55. Thus, the sake review posts were not affected by the number of times the users activated the system. Additionally, we confirmed how each recommendation algorithm affected user behavior on the recommendation system in terms of the popularity of the sake brands. For sake brands registered in the sake-to-drink list, more than half of the sake brands were ranked top 10 of the popularity ranking for all recommendation systems with each algorithm. For the transition to the sake brand description page, the popular sake brands were included to a certain degree for all recommendation systems with each algorithm. Also, the ones with content-based filtering and matrix factorization included sake brands with low popularity rankings (20th or lower) at high ratios of 0.43 and 0.61, respectively. For the sake review posts, all recommendation systems with each algorithm included popular sake brands with rankings of 10 or higher at ratios of more than 0.25. Moreover, the ones with content-based filtering, association analysis, item-based collaborative filtering, and matrix factorization contained popular sake brands ranked 20th or lower at high ratios above 0.46.

5 Discussions

As discussed in the evaluation results in Section 4, a certain number of users actively used the system for posting reviews to each sake brand for each recommendation system, resulting in differences in the numbers of sake brands registered in the sake-to-drink list and transitions to the sake description page, which indicates the influence of the recommendation algorithm. Considering the registered users in the sake-to-drink list as an indicator of items of interest to each user, recommendation systems with association analysis and item-based collaborative filtering showed more effective results. However, considering transitions to the sake description page as an indicator for unknown items to each user, the recommendation systems with content-based filtering, relevance analysis, and matrix factorization showed effective results. In a recommendation system that provides high value-added services, it is required to recommend unknown items of interest to each user. Therefore, association analysis is a useful algorithm for an effective recommendation system, since it has a high index of unknown and interesting items.

Next, we discussed factors contributing to these results from previous studies. We assumed that the recommendation targets and content of the training data attributed to these results. Generally, the interest of users in sake here is influenced by personal preferences and general popularity. In the previous study, it was reported for similar target types, such as movies and books [10], item-based collaborative filtering was best, followed by association analysis [10,12]. However, content-based filtering was reported to be better for targets with strong individual preferences, such as games [9]. Thus, content-based filtering was considered an inappropriate algorithm for the target in this study.

Also, the training data used in the recommendation algorithm can influence the accuracy of the recommendation depending on the period it is collected. In a survey for the dataset of recommending movies and books [10], association analysis, matrix factorization, and item-based collaborative filtering required the latest data from the most recent period to obtain valid results for the recommendation. Since the platforms used in this study are relatively recent ones, the training data used were recently collected data. It is inferred that the recommendation system with association analysis was considered the most effective in the investigation. When using training data collected over a long period, item-based collaborative filtering can yield effective results for the recommendation system. However, this is a topic for future research, as it requires a new investigation. Therefore, under this study, results show that the recommendation system with association analysis was effective in terms of recommendation target and training data.

6 Conclusion

This paper described a practical investigation of how different recommendation algorithms affected the online behavior of users using an online service for sake. To investigate this, we assigned one of four recommendation algorithms to 858 users who logged in during 80 days. The numbers of system logins, transitions to the description page of each sake brand, sake brands on the sake-to-drink list, and posted review comments of each sake brand were analyzed for each recommendation algorithm from the system history. Results show that the recommendation system with association analysis was useful as a sake brand recommendation system, considering individual preference and popularity of the sake review sites. Empirical studies of such a recommendation system are few since it can be conducted only on the service provider side. However, it is useful knowledge for designing a commercial recommendation system, and we believe the findings of this study have contributed in that light.

In our future work, we would like to analyze the influence of the recommendation algorithm on the content of the posted reviews. In this study, the posted review data were collected compared, but qualitative differences between them were not confirmed. However, there may be some influence on the content of the posted reviews since we observed that different algorithms influenced recommended items and user behavior. Thus, it is necessary to use support techniques, such as natural language processing and text mining, but this is not covered in this study because of its large scale of analysis. Additionally, we will investigate the effectiveness of the recommendation system for a long period. Since the training data for a recommendation algorithm is affected by the length of period for which the training data are collected, the finding from the new investigation will contribute to the design of a commercial recommendation system.

References

[1] T. H. Silva, P. O. de Melo, J. Almeida, M. Musolesi, and A. Loureiro, “You are what you eat (and drink): Identifying cultural boundaries by analyzing food & drink habits in foursquare,” In Proc. the Eighth International AAAI Conference on Weblogs and Social Media, .465-475, pp2014.

[2] L. T. Wright, C. Nancarrow, and P. M. Kwok, “Food taste preferences and cultural influences on consumption,” British Food Journal, Vol.103(5), pp.348-357, 2001.

[3] S. Higgs and J. Thomas, “Social influences on eating,” Current Opinion in Behavioral Sciences, Vol.9, pp.1-6, 2016.

[4] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative filtering recommender systems,” In: P. Brusilovsky, A. Kobsa and W. Nejdl, Eds., The Adaptive Web, LNCS 4321, Springer-Verlag, Berlin Heidelberg, pp.291-324, 2007.

[5] M. Deshpande, and G. Karypis, “Item-Based Top-N Recommendation Algorithms,” ACM Transactions on Information Systems, Vol.22(1), pp.143-177, 2004.

[6] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, “Analysis of recommendation algorithms for e-commerce,” In Procs the 2nd ACM Conference on Electronic Commerce, pp.158-167, 2000.

[7] N. Kotonya , P. D. Cristofaro, and E. D. Cristofaro, “Of Wines and Reviews: Measuring and Modeling the Vivino Wine Social Network,” In Procs. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp.28-31, 2018.

[8] J. Y. Seo, , S. W. Han, and H. M. Lee, “A Wine Recommendation System using Collaborative Filtering,” Advanced Science Letters, Vol.23(10), pp.10394-10398, 2017.

[9] D. Jannach and K. Hegelich, “A case study on the effectiveness of recommendations in the mobile internet,” In Procs the 3rd ACM conference on Recommender systems,pp.205-208, 2009.

[10] Z. Huang, D. Zeng, and H. Chen, “A comparative study of recommendation algorithms in e-commerce applications,” IEEE Intelligent Systems, Vol.22(5), pp.68-78, 2007.

[11] B. Pradel, S. Sean, J. Delporte, S. Guerif, C. Rouveirol, N. Usunier, F. FogelmanSoulie, and F. Dufau-Joel, “A case study in a recommender system based on purchase data,” In Procs. the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp.377-385, 2011.

[12] I. Benouaret and S. Amer-Yahia, “A Comparative Evaluation of Top-N Recommendation Algorithms: Case Study with Total Customers,” In Procs. 2020 IEEE International Conference on Big Data, pp.4499-4508, 2020.

[13] M. J. Pazzani and D. Billsus, “Content-based recommendation systems,” In: Brusilovsky P., Kobsa A., Nejdl W. (eds) The Adaptive Web. Lecture Notes in Computer Science, Vol.4321. Springer, Berlin, Heidelberg, pp.325-341, 2007.

[14] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," IEEE Internet Computing, Vol.7(1), pp.76-80, 2003.

[15] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," In Procs. the 20th International Conference on Very Large Data Bases, pp.487-499, 1994.