J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

Dessert Ateliers Recommendation Methods for Dessert E-commerce Services☆

손 연 빈[1] 장 태 우[1] 최 예 림[1*]
Yeonbin Son Tai-Woo Chang Yerim Choi

ABSTRACT

Dessert Ateliers (DA) are small shops that sell high-end homemade desserts such as macaroons, cakes, and cookies, and their popularity is increasing according to the emergence of small luxury trends. Even though each DA sells the same kinds of desserts, they are differentiated by the personality of their pastry chef; thus, there is a need to purchase desserts online that customers cannot see and purchase offline, and thus dessert e-commerce has emerged. However, it is impossible for customers to identify all the information of each DA and clearly understand customers' preferences when buying desserts through the dessert e-commerce. When a dessert e-commerce service provides a DA recommendation service, customers can reduce the time they hesitate before making a decision. Therefore, this paper proposes two kinds of DA recommendation method: a clustering-based recommendation method that calculates the similarity between customers' content and DAs and a dynamic weighting-based recommendation method that trains the importance of decision factors considering customer preferences. Various experiments were conducted using a real-world dataset to evaluate the performance of the proposed methods and it showed satisfactory results.

☞ keyword : dessert atelier, dessert e-commerce, dessert atelier recommendation, clustering, dynamic weighting

1. Introduction*

Recently, the public interest in Dessert Ateliers (DA), small shops selling high-end homemade desserts, has been rapidly increasing in South Korea for the following reasons. In order to relieve stress from daily life, there is a trend that people are pursuing small but sure happiness. By this phenomenon, new words such as "YOLO" and "FORME" have occurred, which stands for "You Only Leave Once" and "For health, One, Recreation, More convenient, Expensive", respectively. Also, for example, the dessert nomad family, who seeks happiness in everyday life by demanding high quality homemade desserts and uploads pictures of them to social network services (SNS) rather than simply eating snacks to fill their stomachs, has emerged [1]. As this proves, the number of SNS posts tagged with "dessert" was 5,100,000 and those tagged with "dessertgram" was about 840,000 [2].

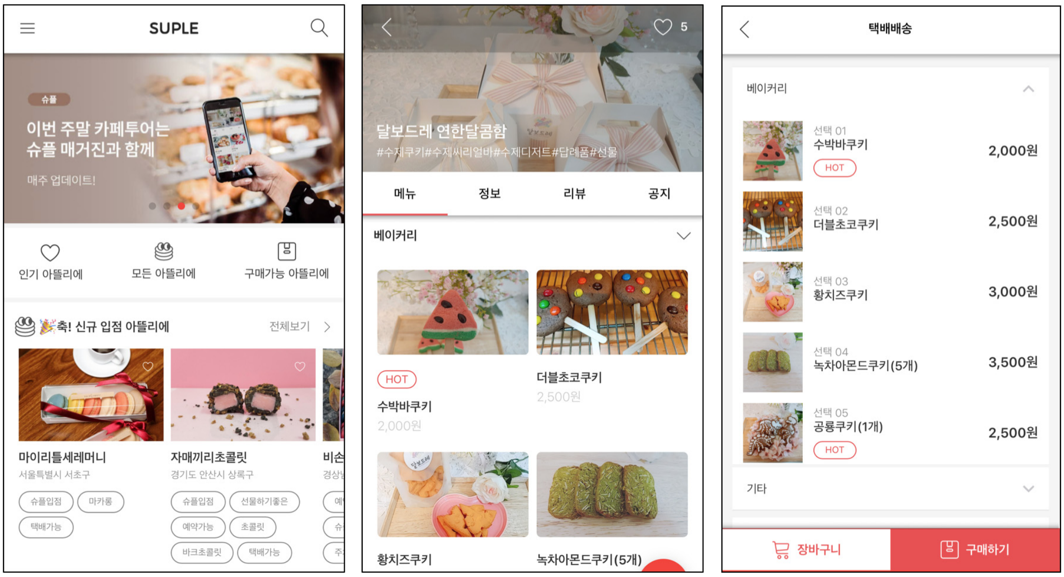
(Figure 1) Dessert photographs that use the same ingredients but have their own characteristics

DA sells desserts that reflect the individuality of the pastry chefs of each DA, so even the same kinds of desserts are slightly different in taste and appearance. Normally, DA has a small space and the number of desserts will be limited because a pastry chef makes the desserts and operates the atelier simultaneously. The personality of the pastry chef is reflected in their desserts, thus even the same kind of dessert may not have exactly the same taste, shape, and so on. For example, Figure 1 shows the difference in visual characteristics for different DAs even in the same menu.

Depending on the characteristic of this dessert atelier, dessert e-commerce has emerged that allows customers to

1 Department of Industrial and Management Engineering, Kyonggi University, Suwon, Republic of Korea
* Corresponding author (yrchoi@kgu.ac.kr)
[Received 1 August 2019, Reviewed 8 August 2019(R2 28 October 2019), Accepted 7 November 2019]
☆ This work was supported by Kyonggi University's Graduate Research Assistantship 2018.
☆ A preliminary version of this paper was presented at ICONI 2018.

http://dx.doi.org/10.7472/jksii.2020.21.1.111



(a) (b) (c)

(Figure 2) Screenshots of dessert e-commerce 'SUPLE'(a) main page of the application (b) dessert atelier detail page (c) dessert purchase page

purchase desserts online for delivery to fulfill the needs to purchase unique desserts sold at remote ateliers [3], and it is represented in Figure 2. Dessert e-commerce provides information such as location, menu, price, and pictures of DA in the country, and can buy desserts from DAs online. It is impossible to purchase the same menu items every day because there is daily variation in quantity and menu because of the ingredients and hand-made factor. However, in dessert e-commerce, it is possible to choose the dessert immediately and conveniently.

When a dessert e-commerce recommends a DA that customers will like based on their past preferences or content, it can improve customer demand. Choosing a DA from which to purchase a dessert is difficult, unlike purchasing ready-made items. When purchasing ready-made items such as a laptop, customers simply filter the conditions like manufacturer, model name, color, and price. However, because the characteristics of desserts cannot be defined by these conditions, simple filtering is not possible.

Although e-commerce, which sells products in various domains, introduced a recommendation system, sales volume increased, but no research on the recommendation system for dessert e-commerce has been conducted. Netflix, a video subscription service, introduced a recommendation system based on the customer's video preference [4], and Amazon, a book e-commerce, introduced a recommendation system based on the customer's past purchase history [5]. However, since dessert e-commerce has recently presented, no research on the recommendation system considering the characteristics of the dessert has been conducted.

Therefore, in this paper, various DA recommendation methods for dessert e-commerce are proposed such as clustering-based DA recommendation and dynamic weighting-based DA recommendation. The remainder of this paper is organized as follows: Section 2introduces dessert atelier recommendation methods; In Section 3, we validate the effectiveness of the proposed methods by an experimental evaluation on areal-world dataset; Section 4 outlines the conclusions.

2. Dessert Atelier Recommendation Methods

2.1 Overview

In this step, two kinds of recommendation method which is available for recommendation to a DA in dessert e-commerce, clustering based DA recommendation, and dynamic weighting (DW) based DA recommendation. The clustering based DA recommendation method forms clusters based on the content similarity between the customers and the DAs, and recommends that the same cluster is regarded as having similar properties. DW based DA recommendation learns the importance of the various determinants that are considered and decides on DA based on the customer's past preferences.

2.2 Clustering-based Dessert Atelier Recommendation

Clustering methods are applied to group similar objects on the basis of their characteristics into a group called a cluster. The purpose of clustering is to maximize the inner similarity within a cluster and minimize the similarity between different groups [6]. Many previous studies have employed the clustering method to recommend food [7]. In this paper, the proposed method executes k-means clustering to cluster the customers and the DAs.

2.2.1 k-means Clustering and Recommendation

k-means clustering is a popular clustering method that requires the input parameter, k, and clusters participants into k groups. The process of k means clustering is as follows. First, the k centroids that represent the mean value of each cluster are randomly picked from the whole dataset. Next, the remaining participants are assigned to the nearest cluster based on the distance between each participant and the centroid of the cluster. This means that the new mean value of the cluster is calculated and is updated to the centroid. This process is iterated until there is no change from the centroid. k-means clustering is executed by the representation vector of customers and DAs.

As a result, we can obtain k clusters with customers and ateliers together, recommending DAs in the same cluster as customers. When the customer and the DA exist in the same cluster, it means that they have similar characteristics. Therefore, it is possible to recommend that DA to the customer.

2.3 Dynamic Weighting-based Dessert Atelier Recommendation

This part describes the details of the dessert atelier recommendation method, with particular consideration of the weight of each visit decision factor. The customer considers various factors in the process of purchasing dessert, which is called a visit decision factor. When we know which of these factors are more important than others, it is helpful to recommend a dessert, and it is called each factor's Preference Sensitivity (PS). PS is calculated based on the deviation of the preference sensitivity from the average of preference, the PS is utilized to calculate the weight of the visit decision factors. This method has been actively studied in the field of cooking recipe recommendations [8].

2.3.1 Average Preference Calculation

We denote the set of customers by $C = \{c_i | i = 1, ..., I\}$, where I is the number of customers in this problem. The set of ateliers is denoted by $A = \{a_j \mid j = 1, \cdots, J\}$, where J is the number of ateliers. The customers' average preference (AP) is calculated using the rating matrix R and the visit decision factor matrix V. R is constructed with ratings given by c_i to a_j and it is denoted by r_{ij}. Moreover, V is constructed with the value of the visit decision factor v_{jk}, where the value of a_j by the $k-$th factor, v_k.The row of V represents the ateliers and the column of V represents the factors. The AP by the factors is calculated by multiplying the ratings and the values of the factors as is shown in Equation 1:

$$AP_k = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ik} v_{jk}}{I} \quad (1)$$

2.3.2 Weighted Preference Score Calculation

The atelier is recommended based on the rank, which is obtained by calculating the weighted preference score of c_i. The weight of v_k by c_i and w_{ik} is calculated from the preference sensitivity S_{ik} between AP_k and the ratings from c_i to v_k, r_{ik}; it is calculated as follows:

$$S_{ik} = \frac{|\ r_{ik} - AP_k\ |}{\min_{l \in \{K_l, K_r\}} |l - AP_k|} \quad (2)$$

$$r_{ik} = \frac{\sum_{i \in I} r_i v_k}{I} \quad (3)$$

$$w_{ik} = \frac{S_{ik}}{\sum_{k \in v_k} S_{ik}} \quad (4)$$

The weighted preference score for a_j and c_i is calculated by multiplying v_k with w_{ik} as is shown in Equation 5:

$$PS_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{J} w_{ik} v_{jk}. \quad (5)$$

3. Experiment

In this section, experiments were conducted to evaluate whether the methods introduced in the previous section performed well. Data was collected from the real dessert e-commerce and used for the experiment. In order to evaluate the performance, experiments were conducted on various evaluation metrics using comparative methods.

3.1 Dataset

Experimental data for executing the clustering-based recommendation method was obtained from a dessert e-commerce application named “SUPLE.” Table 1 presents detailed statistics of the data for each proposed method. The total number of customers is 20 and the total number of ateliers is 200. The factors that were used in each proposed method are divided into two datasets. The dataset for the clustering-based recommendation method consists of keywords and menu descriptions of ateliers. The dataset for dynamic weighting-based RS consists of the number of open days, menus, the number of SNS followers, and the average price of desserts.

(Table 1) Summary of the dataset for each methods, especially the factors

Method	Number	Factors
Clustering based	4	Keywords, DA description, menu names, menu description
DW based	4	The number of open days, menu name, the number of SNS followers, the average price of desserts

3.2 Settings

There are various evaluation metrics for assessing the performance of the proposed recommendation methods; these metrics are divided into two categories: decision-support accuracy metrics and statistical accuracy metrics [1]. This paper uses inner similarity for the clustering-based recommendation method as the decision-support accuracy and mean absolute error(MAE) as the statistical accuracy metrics for the dynamic weighting-based recommendation method. Inner similarity is a measure that calculates the similarity of objects that construct a cluster [9]. MAE is a well-known method for evaluating the performance of a method in the field of recommendations [10].

The cosine similarity was calculated from pairing objects in the same cluster as the inner similarity of a cluster. The larger the value of the similarity, the more evidence that the clustering is performed well. In this paper, we compare the inner similarity values according to k, which represents the number of clusters.

MAE was calculated by comparing the predicted rank of DAs in the proposed DW-based recommendation method and the actual rank that was surveyed from customers; it is shown in Equation 6

$$MAE = \frac{\sum_{n} |r_a - r_p|}{n}. \tag{6}$$

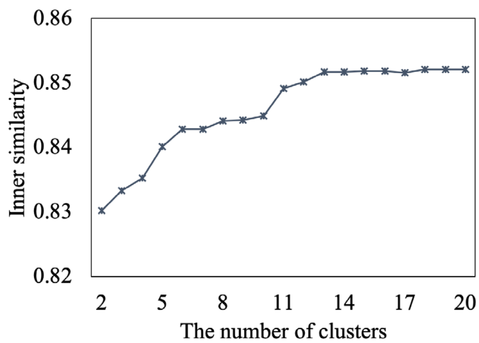
r_a represents the actual rank of DAs that were surveyed by customers, r_p means the predicted rank of a DA based on the proposed method, and n means the number of test data. To validate the performance of the DW-based recommendation method, the actual rank was compared with two kinds of predicted rank: the predicted rank based on the matrix factorization (MF) method, which is a well-known method in the field of recommendation methods and the DW-based recommendation method. MF is the method that factorizes the rating matrix into low-rank matrices [11, 12]. The dimension of factorization needs to be set during the execution process of MF, and this variable was set to 100.

3.3 Experimental results

In order to evaluate the performance of the clustering-based recommendation method, inner similarity values for each cluster were calculated, and it is shown in Figure 3. First of all, the value of inner similarity increases as the number of clusters increases, and larger inner similarity means that instances within the same cluster are similar. Since the value of k is 13, the value of the inner similarity has almost not increased, and it can be said that the optimal number of k is 13. The average similarity was 0.84, which is relatively large. Moreover, the satisfaction rate of the recommended DA for the customers is 87%, which is relatively high.

However, this method shows that similarity is calculated highly even if there is a similar word in the content representing DA but with a completely different meaning. In a dataset, among the DAs used as test subjects, the third and fifth DAs both sell a strawberry-flavored dessert, but third DA uses strawberry artificial flavor and fifth DA uses fresh strawberry. Both DAs sell strawberry-flavored desserts, so the similarity may be high, but the use of artificial flavors and fresh fruits is very different in desserts. Therefore, it is necessary to figure out the introduction of additional methods that can consider these problems.

The DW-based recommendation method shows superior performance compared with the comparative method, and it is represented in Table 2. In order to evaluate the performance of the proposed method, MF, widely used as the recommend method, was used as a comparative method. The results are described as two types: MAPE, which can take into account the ratio of MAE and error fluctuation, and MAE which is the metric using the absolute value difference. From the comparison, the absolute error in applying the MF-based recommendation is larger than the DW-based, and the variation range of error is also wider when MF-based recommendation was applied. In the other words, the proposed DW-based recommendation method outperformed the compared method, MF; this result implies that considering the preference consistency between customers and DAs is effective.



(Figure 3) Performance of the clustering-based recommendation method

(Table 2) Performance ofthe DW based recommendation method

	MF	DW
MAE	2.01	1.83
MAPE (%)	0.38	0.27

4. Conclusion

We have proposed efficient methods for carrying out DA recommendations in a dessert e-commerce system. The key novelty of our approach is as follows: (1) we use a dataset that is usually held in thee-commerce service, (2) the proposed methods can be easily applied to services. This allows customers to reduce the time they spend hesitating and maximize their satisfaction with their purchase. The proposed methods result in substantially better performance on a wide variety of real-world data sets.

The clustering-based recommended method of measuring similarity performed well, however, this method was unable to detect similar word differences. For this reason, a clustering-based recommendation method with a trust similarity measure can be proposed to improve the pre-mentioned shortage. In future developments, various performance measures and methods will suggest for use in dessert e-commerce.

참고문헌(Reference)

[1] Gong, S., “A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering”, Journal of Software, Vol.5, No.7, pp.745-752, 2010.
http://doi.org/10.4304/jsw.5.7.745-752

[2] Instagram, https://instagram.com/

[3] SUPLE (Desserting Inc.), https://suple.me

[4] Gomez-Uribe, C.A., & Hunt, N., “The Netflix Recommender System: Algorithms, Business Value, and Innovation”, ACM Transactions on Management Information System, Vol.6, No.4, pp.13-19, 2016.
http://doi.org/10.1145/2843948

[5] Linden, G., Smith, B., & York, J., “Amazon.com Recommendations: Item-to-item Collaborative Filtering”, Institute of Electrical and Electronics Engineers Internet Computing, Vol.1, pp.76-80, 2003.
http://doi.org/10.1109/MIC.2003.1167344

[6] Kanungo, T., Mount, D. M., Netanyahu, N. S.,Piatko, C. D., Silverman, R., & Wu, A. Y., “An Efficient k-means Clustering Algorithm: Analysis and Implementation”, IEEE Transactions on Pattern Analysis& Machine Intelligence, Vol.7, pp.881-892, 2002. http://doi.org/10.1109/TPAMI.2002.1017616

[7] Phanich, M., Pholkul, P., and Phimoltares, S., “Food Recommendation System Using Clustering Analysis for Diabetic Patients”, 2010 International Conference on Information Science and Applications, pp.1-8, 2010.
http://doi.org/10.1109/ICISA.2010.5480416

[8] Tao, L., Cao, J., & Liu, F., “Dynamic Feature Weighting Based on User Preference Sensitivity for Recommender Systems”, Knowledge-Based Systems, Vol.149, pp.61-75, 2018.
https://doi.org/10.1016/j.knosys.2018.02.019

[9] Ghani, I., “An Improved Combined Content-similarity Approach for Optimizing Web Query Disambiguation”, Journal of Internet Computing and Services, Vol.16, No.06,pp.79-88, 2015.
https://doi.org/10.7472/jksii.2015.16.6.79

[10] Y. Son, H. Ahn, Y. Choi, “A Proofreader Matching Method Based on Topic Modeling Using the Importance of Documents”, Journal of Internet Computing and Services, Vol 19, No.4, pp.27-33, 2018.
https://doi.org/10.7472/jksii.2018.19.4.27

[11] Koren, Y., Bell, R., & Volinsky, C., “Matrix Factorization Techniques for Recommender Systems.”, Computer, Vol.42, No.8, 2009.
https://doi.org/10.1109/mc.2009.263

[12] Baltrunas, L., Ludwig, B., & Ricci, F., “Matrix Factorization Techniques for Context Aware Recommendation.”, Proceedings of the fifth ACM conference on Recommender systems., pp.301-304, 2011.
https://doi.org/10.1145/2043932.2043988

◐ 저 자 소 개 ◑

손 연 빈(Yeonbin Son)
2018년 경기대학교 산업경영학과(공학사)
2017년~현재 Desserting Inc. (SUPLE) 대표
2018년~현재 경기대학교 산업경영공학과 석사과정
관심분야 : 머신러닝, 추천시스템
E-mail: yeonbin517@gmail.com

장 태 우(Tai-Woo Chang)
1995년 서울대학교 산업공학과(공학사)
1997년 서울대학교 산업공학과(공학석사)
2004년 서울대학교 산업공학과(공학박사)
2002년~2007년 한국전자통신연구원 연구원/선임연구원
2007년~현재 경기대학교 산업경영공학과 교수
2017년~현재 경기대학교 지능정보융합제조연구센터(GRRC) 센터장
관심분야 : 스마트공장, 지능정보기술, 시스템분석
E-mail: keenbee@kgu.ac.kr

최 예 림(Yerim Choi)
2010년 서울대학교 산업공학과(공학사)
2016년 서울대학교 산업공학과(공학박사)
2016년~2017년 네이버랩스 Data Scientist
2017년~현재 경기대학교 산업경영공학과 조교수
2019년~현재 ai.m Inc. 대표이사
2019년~현재 경기대-KT 빅데이터센터 부센터장
관심분야 : 인공지능/머신러닝, 빅데이터 기반의 인간 모델링
E-mail: yrchoi@kgu.ac.kr