A sequence-based personalized service for the short life cycle products

Ju-Choel Choi
Dept. of Business Incubator Kyung Hee University

수명주기가 짧은 상품들에 대한 시퀀스 기반 개인화 서비스

최주철
경희대학교 창업보육센터

Abstract Most new products not only suddenly disappear in the market but also quickly cannibalize older products. Under such a circumstance, retailers may have too much stock, and customers may be faced with difficulties discovering products suitable to their preferences among short life cycle products. To address these problems, recommender systems are good solutions. However, most previous recommender systems had difficulty in reflecting changes in customer preferences because the systems employ static customer preferences. In this paper, we propose a recommendation methodology that considers dynamic customer preferences. The proposed methodology consists of dynamic customer profile creation, neighborhood formation, and recommendation list generation. For the experiments, we employ a mobile image transaction dataset that has a short product life cycle. Our experimental results demonstrate that the proposed methodology has a higher quality of recommendation than a typical collaborative filtering-based system. From these results, we conclude that the proposed methodology is effective under conditions where most new products have short life cycles. The proposed methodology need to be verified in the physical environment at a future time.

Key Words: Short life cycle product; Personalized service, Sequence, Preferences, Recommender system

요. 거의 대부분의 신상품들은 시장에서 급격히 사라질 뿐만 아니라 기존 상품들의 매출감소를 불러온다. 이처럼 수명주기가 짧은 상품으로 인해 소비자들은 과다한 재고를 보유하게 될 뿐만 아니라 소비자들은 자신의 선호를 반영하는 제품들을 발전하는데 어려움을 겪는다. 이러한 문제를 해결하기 위하여 본 연구에서는 시장에 따라 변화하는 소비자의 선호를 반영한 추천 방법론을 제안하였다. 제안한 방법론은 소비자의 동적 선호 유형 파악, 네이버, 모바일 이미지 거래 데이터를 이용하여 제안된 방법론의 유효성을 검증하였다. 실험결과 제안된 방법론의 추천 정확도가 전통적인 협업필터링의 정확도보다 높았다. 이러한 결과를 통해, 본 연구에서 제안한 방법론이 짧은 수명주기를 가진 제품을 추천하는 데 효과적이라는 결론을 내릴 수 있다. 따라서 향후 제안된 방법론을 현업에 적용하여 실제적 유효성을 검증할 필요가 있다.

주제어: 짧은 생명주기 상품, 개인화서비스, 시퀀스, 선호도, 추천 시스템

Received 2 November 2017, Revised 1 December 2017
Accepted 20 December 2017, Published 28 December 2017
Corresponding Author: Ju-Choel Choi
(Dept. of Business Incubator Kyung Hee University)
Email: chojc@khu.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Digital Convergence  |  293
1. Introduction

Currently, many types of new products are launched on the market. Most new products not only disappear in the market as times passes but also quickly cannibalize older products. According to the Nielsen report[1], not only do many new products expire within the first three years after their launches but also approximately 50% of customers are pleased to switch to a new brand. Therefore, the fast change in customer preferences for the product makes the product life cycle shorter.

Short life cycle products cause demand forecasting to be much more difficult for retailers. Consequently, retailers may keep excessive stocks to prevent stockouts. Moreover, customers may experience difficulties discovering products suitable to their preferences among short life cycle products because a large number of short life cycle products such as electronic products, multimedia contents, seasonal products, and fashion products appear suddenly and disappear quickly[2]. To address these issues, many retailers such as Amazon[3] and Netflix[4] introduced recommender systems, which help customers to find products suitable to their preference by aggregating and analyzing other customers' suggestions such as product purchases or reviews[5, 6]. These systems have brought many tangible and intangible benefits such as cross-selling opportunities, the enhancement of customer royalty, and stock portfolio management to retailers[7, 8, 9]. In particular, collaborative filtering (CF) is one of the most successful techniques in the recommender system[9, 10, 11, 12].

CF is a recommendation technique that utilizes a historical dataset such as retail transactions and web-logs to search the k nearest neighbors, who have a high degree of similarity with the customers[5, 8, 10, 14]. Such a typical CF-based recommender system has potentially posed some problems such as sparsity problems, cold-start problems, and scalability problems[13, 14, 15]. Therefore, most studies have focused on discovering solutions to these problems and have compared the accuracy of the proposed methodologies against that of the benchmark system, i.e., a typical CF-based recommender system[16, 17, 18, 19, 20, 21, 22].

However, most products within a historical dataset have a short life cycle. Furthermore, customers in general tend to purchase the latest version of products instead of old products [23,24]. Therefore, it is important to discover customer-specific behavior patterns under such conditions. Because the shorter is the product life cycle, the less probable would customers discover products suitable to their preferences.

One type of historical dataset is the time-series dataset, which is composed of customer information, product information, and sequential order information between products[25]. Particularly, sequential order information contains changes of a customer's preferences regarding the products over time[26, 27, 28]. Therefore, discovering customers who have similar purchasing sequences could help to enhance the quality of recommendations on the short life cycle products because most customers do not prefer outdated products.

Therefore, we propose a recommendation methodology for short life cycle products. In general, one of the key phases in the typical CF-based recommender system is creating a customer profile. Most previous studies created a customer profile in the form of a customer-product matrix. However, they do not consider customers' purchasing sequences. To reflect changes in customer preferences regarding short life cycle products, we create a customer profile in the form of a customer-sequential elements matrix. Then, we find neighbors of a target customer and generate a recommendation list for him/her.
2. Related work

2.1 Short life cycle products

In the fashion-driven market, companies constantly launch new products or services to maintain a competitive advantage. This causes short product life cycles. For example, Apple and Samsung launch new flagship smartphones such as an iPhone model and a Galaxy S series model to solidify their position in the competition in the smartphone market every year. However, sales of their products decline quickly when new models of the smartphones are introduced. That is, the short product life cycle leads to obsolescence costs[29]. Thus, it is important to forecast demand and manage inventory of a short life cycle product.

Previous studies related to a short life cycle product are broadly classified into two major classes: demand forecasting and inventory management. First, many models are applied for the forecast demand of a short life cycle product. Kurawarwala & Matsuz[30] studied a linear growth model, exponential growth model, and seasonal trend growth model based on life cycle growth and seasonality to forecast demand in the PC industry. Xu & Song[31] applied a bass model to forecast demand of a short life cycle product. Furthermore, Trappey & Wu [32] forecasted demand of a short life cycle based on the time-varying extended logistic, simple logistic, and Gompertz models. Second, most studies on inventory management are related to supply chain management. Higuchi & Trout [33] proposed a simulation model of the supply chain for a short life cycle product such as Tamagotchi. They argued that it is important to set supply chain specifications in advance to improve the benefits of a short life cycle product. Liu et al [34] have studied a selling system to address coordination with the pricing contract and the ordering-quantity contract in the IT industry.

The object of our study is to forecast demand of a specific product that has a short life cycle at the micro-level and to recommend the specific product to a target customer. In other words, our object is to develop a recommendation methodology for forecasting the purchase likelihood of a specific product.

2.2 Sequence-based recommender system

A sequential dataset comprises sequences of the ordered events such as purchase and web access. Such a sequential dataset is useful for discovering a customer's behavior pattern[35, 27, 28]. There are many technologies involving analysis of a sequential dataset. The most typical technology is sequential pattern mining. The representative approaches for sequential pattern mining are GSP (generalized sequential patterns), SPADE (sequential pattern discovery using equivalent classes), and PrefixSpan (prefix-projected sequential pattern mining). However, sequential pattern mining cannot perform personalized recommendations because sequential pattern mining does not offer customer-specific sequential pattern[35].

Therefore, some researchers have developed recommendation methodologies considering sequences. Cho et al.[36] proposed a hybrid methodology that combined data mining techniques, such as self-organizing map and association rule mining, and CF. Hamir et al.[36] suggested a new methodology based on context information for music recommendations. In particular, sequential pattern mining has been applied for capturing changes in contextual states. Moon et al[37] have obtained visitors' locations and booth visit timestamps through ubiquitous technologies and have proposed a recommendation methodology considering booth visit sequences in off-line exhibition environment. Salehi et al.[38] developed a material recommender system for learners in the e-learning environment. They have applied two sequence mining algorithms, including the Apriori algorithm, PrefixSpan algorithm, and multidimensional CF, to the recommender system. However, most previous studies have rule consistencies such as rule redundancy and rule conflict.
because they generate rule sets.

Therefore, we propose a memory-based methodology to discover user-specific behavior patterns and recommend a short life cycle product to a customer. The proposed methodology follows the principle of CF although the CF is modified.

3. Methodology

3.1 overall procedure

Customer preferences regarding products change over time. Nevertheless, most previous studies on CF have regarded customers who purchased similar product sets as neighbors without consideration of changes in customers' preferences. However, we assume that one customer who purchased a product a after purchasing a product b is not similar to the other customer who purchased a product b after purchasing a product a.

Our proposed procedure basically follows the principle of CF. However, we propose a new creation method of a customer profile considering changes in customers' preferences over time. Our recommendation procedure consists of the following three phases. In the first phase, we create a customer profile considering a customer's purchasing orders. In the second phase, we discover customers as known neighbors through a similarity measure. In the final phase, we generate recommendation lists of what a target customer is likely to purchase.

3.2 Step 1: customer profile creation

Given a historical dataset of m customers on n products, a typical CF-based recommender system represents a customer profile as the form of m x n customer-product matrix, R, such that r_{ij} is one if the $i^{th}$ customer purchased the $j^{th}$ product, and zero, otherwise [21]. However, its method to create a customer profile causes information loss because a historical dataset is a time-series dataset comprising sequences of purchasing events.

Given a historical dataset of m customers on n products consisted of a set of tuples of the form <customer_ID, product_ID, timestamp>, the proposed customer profile is created in four steps as follows. First, we extract an initial customer profile in the form of $m \times n$ customer-product matrix such as a typical CF-based recommender system. The initial customer profile, $R=(n_j)$, is defined as follows:

$$r_{ij} = \begin{cases} 1, & \text{if the } i^{th} \text{ user purchased the } j^{th} \text{ product} \\ 0, & \text{otherwise} \end{cases}$$

where $i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$; and $m$, $n$ is the total number of customers and products.

Second, we extract a sequence $s$ by customers, which represents an ordered list of purchasing products. For example, the $i^{th}$ customer's sequence is defined as $s_i = \langle x_1, x_2, x_3, \ldots, x_n \rangle$ where $x_j (j=1, 2, \ldots, n)$ is the $j^{th}$ product and $x_1$ is purchased before $x_2$, which purchase $x_3$, and so on. Third, we discover the elements of sequence $E$ by customers, which is denoted as a pair of products by purchase orders. A sequential element is represented as $x_i x_j$, where $i$ is not equal to $j$. For example, the $i^{th}$ customer's sequential elements of such a pair with $s_i = \langle x_1, x_2, x_3, \ldots, x_n \rangle$ are defined as $E_i = \langle x_1 x_2, x_3, \ldots, x_{n-1} x_n \rangle; \ldots, x_1 x_2 x_3, \ldots, x_{n-2} x_n \rangle$; where the first product in the pair is purchased before the second product. Finally, a customer profile is represented in the form of a customer-sequential element matrix. The customer-sequential element matrix, $R=(r_{jk})$, is defined as follows:

$$r_{jk} = \begin{cases} 1, & \text{if the } i^{th} \text{ user purchased the } j^{th}\text{ sequential element} \\ 0, & \text{otherwise} \end{cases}$$

where $i = 1, 2, \ldots, m$; $k=1, 2, \ldots, n_p$; $m$ is the total number of customers, $n$ is the total number of products, and $n_p$ is the total number of sequential elements.
3.3 Step 2: neighbor formation

After creating a customer profile, we calculate the similarity to form neighbors of a target customer. The similarity between the target customer a and another customer b is computed utilizing the Pearson-r correlation coefficient as follows:

\[
sim(a, b) = \frac{\sum_{k=1}^{n_p_2} (r_{a,k} - \bar{r}_a)(r_{b,k} - \bar{r}_b)}{\sqrt{\sum_{k=1}^{n_p_2} (r_{a,k} - \bar{r}_a)^2 \sum_{k=1}^{n_p_2} (r_{b,k} - \bar{r}_b)^2}}, (3)
\]

where \( n \) is the total number of products, \( n_p_2 \) is the total number of sequential elements, \( r_{a,k} \) and \( r_{b,k} \) are a's and b's ratings on the sequential element \( k \), and \( \bar{r}_a \) and \( \bar{r}_b \) are a's and b's average ratings of all sequential elements.

Then, we select 1 customers with a high degree of similarity between the target customer a and other customers as the target customer a's neighbors.

3.4 Step 3: recommendation list generation

Finally, we generate the top-N recommendation utilizing \( PLS(a, j) \), which denotes the purchase likelihood score of the target customer a for the product j.

\[
PLS(a, j) = \frac{\sum_{i \in X} r_{i,j} \times \sim(a, j)}{\sum_{i \in N} \sim(a, i)} (4)
\]

where \( r_{i,j} \) is neighbor i's rating on the product j within the initial customer profile of the step 1, and \( \sim(a, i) \) is the similarity between the target customer a and his/her neighbor i.

Because the PLS value represents the purchase likelihood for a specific product, the higher is the target customer's PLS value for the product, the greater is the likelihood that he/she will purchase the product. Therefore, we sort products according to their PLS values and select N products with high PLS values as the recommendation set.

4. Experimental evaluation

In this section, we illustrate experimental results that evaluate the accuracy of our proposed methodology. All experiments were performed on a PC running Windows 8.1 with Intel core i7-4500U @ 1.80GHz processor, 8GB RAM, 250 SSD and 1TB HDD. All programs were implemented in SQL Server 2014. In addition, we compared the proposed methodology with a typical CF-based recommender system as a benchmark system.

4.1 Datasets

We employed a historical dataset of 713 customers on 3,596 mobile wallpaper images collected from S telecommunication company in Korea during the period between June 1, 2004 and August 31, 2004. That is, the dataset is customers' purchase records on wallpaper images to decorate their mobile screens. The dataset is composed of the customer serial number, image serial number, and purchase date. The distribution of wallpaper images purchased by release date is shown in [Fig. 1]. Old images were no longer sold. However, most images were sold within the first four months after their launches because the images are mostly of current Korean celebrities. That is, the mobile wallpaper images are very fashion-sensitive. Therefore, mobile wallpaper images have short life cycles.

![Fig. 1] The distribution of the distinct wallpaper images by release date
We divide the dataset into a training set and test set for the experiment. The period between June 1, 2004 and July 31, 2004 is set as the training period, and the period between August 1, 2004 and August 31, 2004 is set as the test period. The training set contains 4,480 transactions of 670 customers for 1,920 images, and the test set contains 5,919 transactions of 596 customers for 2,165 images.

4.2 Evaluation metric

Many researchers have employed recall, precision, and F1 to evaluate the accuracy of the recommender systems\cite{13, 21, 34}. Recall and Precision are defined as the ratio of the number of products in both the test set and the recommendation set to the test set size and to the recommendation set size, respectively. However, the recall values increase when increasing the size of the recommendation set. On the contrary, precision values decrease when increasing the size of the recommendation set. To solve such a problem, we calculate F1. F1 is defined as follows:

\[
F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}
\]

4.3 Experimental results

This section presents a detailed experimental evaluation of the different parameters for the steps of our proposed methodologies and compares the performances of the proposed methodology (SCF) with those of a typical CF-based recommender system (CF). We first determine the optimal sizes of the neighbors and second, select the number of recommended products utilizing them. Then, we analyze the distribution of the recommendation set by release date.

First, to determine the sensitivity of neighbor size, several experiments were performed by varying the number of neighbors and computing the corresponding average of the F1 values. \[\text{Fig. 2}\] presents our experimental results. The recommendation quality of CF and SCF dramatically increased as the neighbor size increased, but after a certain peak, they declined modestly. That is, recommendation quality decreases after an optimal size of the neighbor because the recommendation becomes less personalized by over-fitting as the size of neighbors increases. In our experiment, the average of the F1 values in the CF was the highest at 0.0126 when the neighbor size was 10, and that in the SCF was the highest at 0.0228 when the neighbor size was 14. Thus, we utilized 10 and 14 as our choice of neighbor sizes for CF and SCF, respectively.

\[\text{Fig. 2}\] Impact of neighbor size

\[\text{Fig. 3}\] Comparison of recommendation quality

After we determined the optimal sizes of neighbor for CF and SCF, the various experiments were performed to determine the optimal sizes of the top-N recommendation sets. \[\text{Fig. 3}\] presents our experimental results when the number of neighbors for CF and SCF were 10 and 14, respectively. The F1 values of both CF and SCF increase rapidly. Then, they
A sequence-based personalized service for the short life cycle products

decrease moderately after arriving at certain peaks. The optimal top-N size for CF is 43, and its F1 is 0.0130; the optimal top-N size for SCF is 46, and its F1 is 0.0234. That is, SCF achieves an approximately 79.6% improvement in comparison to CF.

Furthermore, we analyzed the ratio of the number of distinct images to the recommendation set by release date when the recommendation quality of both CF and SCF is the highest. The results are shown in [Fig. 4] Overall, both CF and SCF recommend most of the latest images to customers. However, CF recommends the oldest images slightly more than SCF, although the difference between CF and SCF is imperceptible. The experimental results demonstrate that our proposed methodology has the potential to recommend short life cycle products including fashion-sensitive products.

![Graph](image)

[Fig. 4] The ratio of the number of distinct images by release date to the recommendation set

5. Conclusion

In this paper, we have proposed a new methodology for recommendation of short life cycle products. Existing CF-based recommender systems consider whether a customer purchased a product. However, it is important to consider each individual’s purchasing pattern because short life cycle products appear and disappear suddenly. Therefore, we proposed a recommendation methodology considering purchasing sequences. The proposed methodology follows the principles of CF. In particular, we created a customer profile considering changes in customers’ preferences for products. We compared the proposed methodology with a typical CF-based recommender system for evaluation. Then, we analyzed the ratio of the number of distinct images to the recommendation set by release date. Experiments based on real datasets have shown that the performance of the proposed methodology is better than that of a typical CF-based recommender system. Furthermore, the proposed methodology recommends the latest products better than a typical CF-based recommender system.

However, our study has some limitations. First, the proposed methodology has the data sparsity problem inevitably because a pair of products by purchase orders is very large. nP2 sequential elements theoretically exist. Note n is the total number of products. Second, the dataset spans a relatively short time period. In general, questions of generalization arise for small datasets.

Our future work will obtain long-term datasets and then split the long-term spans into short time spans. In addition, we will apply a sliding window analysis to the proposed methodology for a more reliable recommendation.

REFERENCES

algorithm for categorical sequence data." Information 
changes in customer buying behavior for 
collaborative recommendations." Expert Systems 
[27] H. Wang, and S. Wang, "Mining purchasing 
sequence data for online customer segmentation." 
International Journal of services operations and 
[28] Y. Zhang, and J. Cao, "Personalized recommendation 
based on behavior sequence similarity measures." 
In Behavior and Social Computing, Springer 
[29] P. Childerhouse, and D. Towill, "Engineering 
supply chains to match customer requirements." 
growth models for medium-term forecasting of 
short life cycle products." Technological Forecasting 
and Social Change, Vol. 57, No. 3, pp. 169–196, 
1998.
[31] X. H. Xu, and Q. Z. Song, "Forecasting for 
products with short life cycle based on improved 
BASS model." Industrial Engineering 
the time-varying extended logistic, simple logistic, 
and Gompertz models for forecasting short product 
[33] T. Higuchi, and M. D. Troutt, Life cycle 
management in supply chains: Identifying 
innovations through the case of the VCR. Hershey, 
[34] B. Liu, J. Chen, S. Liu, and R. Zhang, 
"Supply-chain coordination with combined contract 
for a short-life-cycle product." Systems, Man and 
Cybernetics, Part A: Systems and Humans, IEEE 
next-items recommendation via personalized 
sequential pattern mining." In Database Systems 
for Advanced Applications, Springer Berlin 
[36] N. Hariri, B. Mobasher, and R. Burke, 
"Context-aware music recommendation based on 
latenttopic sequential patterns." In Proceedings of 
the sixth ACM conference on Recommender 
[37] H. S. Moon, J. K. Kim, and Y. U. Ryu, "A 
sequence-based filtering method for exhibition 
booth visit recommendations." International Journal 
620–626, 2013.
[38] Salehi, M. Kamalabadi I. N. and Ghoushchi. M. B. 
G, "Personalized recommendation of learning 
material using sequential pattern mining and 
attribute based collaborative filtering." Education 
and Information Technologies, Vol. 19, No. 4, 

최주철(Choi, Ju Choei)

· 1988. 2: Kyunghee University, 
   Mechanical Engineering (BS)
· 2005. 8: Kyung Hee University 
   Graduate School (MS)
· 2009. 2: Kyunghee University 
   Graduate School (Ph.D)
· 2012. 3+: Kyunghee University, 
   Associate Professor
· Interests:Recommendation systems, embedded software, 
data mining
· E-Mail choije@khu.ac.kr

Journal of Digital Convergence | 301