Personalized Recommendation on Discount Coupons

Article · December 2010

DOI: 10.5176/978-981-08-7466-7_kd-18 · Source: OAI

3 authors, including:

Teemu Mutanen
VTT Technical Research Centre of Finland

Sami Nousiainen
VTT Technical Research Centre of Finland

Personalized Recommendation on Discount Coupons

Teemu Mutanen
VTT Technical Research Centre of Finland
Espoo, Finland
teemu.mutanen@vtt.fi

Sami Nousiainen
VTT Technical Research Centre of Finland
Espoo, Finland
sami.nousiainen@vtt.fi

He Liang
East China Normal University
Shanghai, China
lhe@cs.ecnu.edu.cn

***Abstract*—We present several recommendation approaches how discount coupon service could be personalized in order to cut down customer's usage time. The Velo system is already implemented coupon service system in China with millions of users. The proposed approach makes use of customer and dispenser meta data together with previous user coupon prints in order to predict upcoming prints for a customer. The implementation of the service, data collection, and how the implementation affects the customers' behavior all have an effect on the accuracy of the printed coupons. The analysis of customer characteristics help the service to present relevant coupons to a customer in order both to raise customer's satisfaction and to aid customers on finding relevant coupons faster. The results show that item recommendation systems provide valuable benefits for coupon service.**

***Keywords*-Personalization; Discount Coupon; Recommendation Systems; User Modeling;**

I. Introduction

Discount coupons are used to promote products or to differentiate between customers based on price sensitivity. Coupons will also attract new customers and activate old; from business point of view: pull in business. Coupons serve a particular purpose as they can engender market segmentation whereas lowering regular prices cannot [1]. In this paper we present several recommendations for discount coupon personalization.

To engage consumers, an efficient coupon delivery system is important. In China, there are coupon printing machine located in various places where consumers can obtain the coupons themselves. These machines hold tens, sometimes even hundreds of coupons from a wide range of different companies. In this setting, a consumer may have trouble finding the coupons he wants or needs as the most interesting coupon may be buried under tens of other coupons. By personalized filtering of the coupons based on recommendation system can help both consumers and the coupon dispenser service providers.

In this paper we present several recommendations for discount coupon personalization. To the best of our knowledge, this is a novel research topic in the area of recommendation systems. We have identified several prediction use cases related to the coupons, ranging from more general (predict coupon printing amounts per customer or the probability of new customer becoming a loyal customer) to more specific (predict which coupon the customer will print). Here, we are focusing on the more specific questions. The real deployment scenario that we are interested in is: *When the user is in the front of the coupon dispenser (printer), recommend coupons to him/her*. However, in this paper we utilize data already collected from a commercially operational system and thus do not deploy any recommendation algorithms in the system. Therefore, the problem statement for this study can be summarized as: *Predict which coupon the user has printed out of the set of coupons that were available at the dispenser at that instant of time when the user printed a coupon*. In addition, we analyze the value the implicit and explicit information has on the predictions.

This paper makes the following contributions: (1) a problem description on a novel research topic, (2) experiments of different recommendation algorithms in this topic and (3) discuss possible future reserch aspects and implementations.

The paper is organized as follows: in II-A we discuss related research, in II-B we present the coupon dispenser service in more detail and in II-C we discuss the data which has been collected and its deficiencies. We present our data used in recommendations in II-D. In Chapter III we discuss customer characteristics and in Chapter IV we present several recommendation approaches. Results of our recommendations are presented in V and in Chapter VI we state our conclusions and discuss about future research topics.

II. Background

A. Related Work

To the best of our knowledge personalized discount coupon recommendations are novel area in recommendation systems.

On general level, the recommendation task consists of predicting for each user (customer) score for each (available) item, describing the relevance or interest. Usually there are two types of information sources for achieving this: content information and user clicks/transactions. Content based approaches make use of descriptions of events, items and users. Recommendations in content based methods are for example based on word frequencies. Collaborative filtering (CF) methods use only information about users and items on identity level. The CF recommendations are based on information about user's previously purchased items and users who have purchased items in common with her. For example Burke [2] gives a more detailed description to CF and hybrid models based recommendation algorithms.

Recommendation algorithms have been applied in various domains before. Two of the most widely studied domains are item recommendations based on item similarities, e.g. [3] and predicting item ratings based on user ratings on other items. The latter has provided wide range of scientific research, probably because of availability of large datasets (Netflix[1] and Movielens[2]). For recommendation approaches in predicting user ratings, see for example [4] or [5].

In addition, recommendation systems have also been implemented in the internet for dynamic content by Chu et al. [6]. Chu et al. research article recommendations to the users. Another example of personalized recommendation on dynamic content, which is based on logged in users and their web history is presented in [7] where user profiles are generated by the system based on past clicks.

The value of history data has been studied in market research. In [8] Rossi et al. apply econometric models for predicting future delivery of coupons based on purchase history. The optimal value of the discount in coupons has been studied for example by Ben-Zion et al. [9]. For the company perspective the coupons may be used as a part of larger marketing campaign and coupons can be used as a tool for many targets.

B. Coupon Dispenser Service

Velo [3] is a company providing a discount coupon service in China. Customers can use the coupon service for example from the website; customers can get discount coupons in electronic format over the internet and mobile phone; customers can print discount coupon's using physical dispensers. The Velo company provides the coupon service for cities in China: Shanghai, Beijing, Nanjing, SuZhou and WuXi. More than 3 million users are using the coupon service. What makes this service interesting is that most of the customers like to use the coupons in a printed format. Also majority of companies in China only accept coupons in printed format for the record making convenience.

If one wants to use the Velo discount coupon service, she should first get a Velo Card (an RFID-tag card), which can be acquired for a small fee. For the card to work a registration is needed. During the registration customers can provide personal information voluntarily. Customer with a Velo card can print coupons on any on-site dispensers. An example of an on-site dispenser is shown in Figure II-B. Each dispenser has several buttons and under each button a customer can find several coupons. On the dispenser there is also a button named MyVelo button. With this button the customer can print coupons bookmarked beforehand by her in the service website.

[1] http://www.netflixprize.com/

[2] GroupLens at University of Minnesota. http://www.grouplens.org/node/73.

[3] http://velo.com.cn/

Figure 1. On-site Velo coupon dispenser.

For the service provider, reduction of customer's time when printing coupons is critical. During the peak hours customers are waiting in line for their turn in the dispenser. By helping customers to find relevant and interesting coupons faster the utilization of the dispensers will be increased and more customers can be served on one site. In addition, customer satisfaction may increase while unnecessary coupon prints may decrease.

C. Data Collection

In this section some of the characteristics of the collected data are presented. Because of the various difficulties with data and accuracy of the information, we filtered some transaction records away. Filtering criteria and resulting dataset is presented in Chapter II-D and Table I in more detail.

1) Language: One of the main issues with the data collected was Chinese language. Although we had all variable names and the value names of some class variables translated into English, some of the variable values included text description in Chinese. By discarding the information written in Chinese, we lost also the information about the coupon content and dispenser location.

2) Customer Identification: The coupon service system requires customers to identify themselves by RFID-tags before they can print out coupons. We may expect that the customer is always the one who registered the RFID-tag and to whom we should present recommendation. However, this might not be the case. A customer may print several coupons and give some or all of those to her friends. A customer may borrow the RFID-tag to her friends. From data viewpoint, the coupon print records look all the same.

3) Dispensers: The coupon dispensers were implemented with either 12 or 16 buttons. An example of an dispenser is shown in Figure II-B. However, the number of available coupons varies from 40 to 160, which means that customers can find several coupons under each button. We are not able to detect from the data the order of the coupons, mainly because we lack knowledge about coupon providers.

Dispensers are in various types of locations, e.g. shopping centre or train station. The customer print behavior may be different in each location.

4) Coupons: In the service there is different type of coupons present. Some of the coupons, e.g. coffee, may be printed by the same user many times during the day/week. While some coupons, e.g. jeans/mp3-player, might be printed by the user only once or twice, after which similar coupons are of no interest anymore. There may also be coupons which do not offer a discount, but offer product trials for free. From the transaction records we formulated binary values of the printed coupons by the customer.

Because of the free business model [4] the coupon prints are free and customers do not have a cost printing them. Customers may print coupons for example *just in case* and never use them. From the transaction records we cannot detect which coupons are actually used. We are not able to get sufficient feedback about good coupons. The accuracy of coupon recommendations based on transaction records is limited. We cannot actually interact with the customers. There is no way for us to *ask* the customers if the coupons are really useful, to be able to make relevant recommendations. To achieve this sort of metric would require some modifications on the implementation, which is outside the scope of our research.

D. Filtered dataset

For our problem, presented above, we make use of only part of the data and part of the variables collected from a commercially operational service. The dataset analyzed in this paper was collected from time period 1.8.2009-21.2.2010. The dataset has several attributes which were not useful in our prediction task, some of the problems with these attributes were discussed above, e.g. rfid-tag identification records. We filtered the transaction records from the chosen time period and were left with only the relevant records. The filtering criteria were as follows:

- RFID-tag identification records.
- Non-active coupons and customers, i.e. those coupons which are not printed and customer who do not print anything during the time period.
- Welcome coupons were filtered, the coupons which are presented to the customer during sign in.
- MyVelo coupons were filtered, those coupons which have been bookmarked by the user in the service website. These coupons are not on the list of available coupons to the user.

In addition to these we filtered out those transaction records which had coupons that were not available in the dispenser. [5] For evaluating recommendation algorithms, it is necessary for us to know, which coupons were available at certain dispenser at certain instant of time. This is the set of coupons out of which we should attempt to predict the actual coupon printed as accurately as possible.

The filtering process produced three datasets including characteristics of each coupon, customer and dispenser. The description of the variables in these datasets is presented in the Table I. Only a sample of 10000 customers were selected from the service's overall customer database.

III. CUSTOMER CHARACTERISTICS

In the filtered dataset 10000 customers were selected from the overall customer database.[6] The selection was based on the need to pick those customers with lot of background information. In the registration process the customers may provide background information voluntarily.

1) Duration: Because of the free business model of the coupon service, the ratio of trial versus loyal customers could provide valuable insight on print prediction task. In Figure 2 a histogram of customer lifetime is presented. Customer lifetime was calculated simply by counting the days between first and last print actions in our training dataset. In the dataset customer lifetime was on average 71 days, with standard deviation of 59 days.

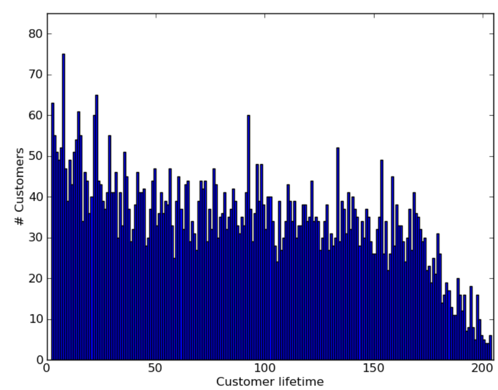

Figure 2. Number of customers based on customer lifetime, i.e. days from the first coupon print record till the last print record in our data set. The customers who only printed coupons one day and never returned (n=1631), where left out of the histogram.

2) Interests: During the registration process customers were also given a possibility to indicate their interest in various topics. [7] The customer marked interest makes it

[4] A business model where customers are given free content and the revenue comes from advertisements or additional paid content.

[5] To the best of our knowledge these coupons might be link coupons or one of the coupons mentioned above. However, these coupons did make only small portion of the total.

[6] From the 10000 customers only 8574 were active customers during our research period.

[7] The interest topics are: food, shopping, cinema, gaming, hairdressing, cosmetics, reading, dancing, body building, club, photo, fashion, outdoors sport, pet, KTV, others

Table I

DESCRIPTION OF THE FILTERED DATASET. ALONG WITH VARIABLE NAMES THE NUMBER OF ACTIVE COUPONS, CUSTOMERS, DISPENSERS, AND PRINT RECORDS WITH TIME PERIOD ARE PRESENTED.

	Customer	Coupon	Dispenser	Print Records
Variables	ID Marriage Company Interest Monthly Industry Post City Area Sex	ID	ID District City	row ID Dispenser ID Customer ID Coupon ID date & time
Details	- # customers:10000 - # active/# non active: 8574/1423	- # coupons: 14832 - # active/# non active: 3141/11691	- # dispensers: 1246 - # buttons in a dispenser 12/16	- time period: 1.8.2009-21.2.2010 - # print records 369760

possible in research how valuable this information is and could personalization be based on it? We produced a binary vector of interest for each customer. The assumption was that only liking could be indicated, a blank answer and dislike are considered the same. We produce two types of similarity recommendations based on these interest vectors, described in IV-C1.

3) Profile: From the customer background information we selected the class attributes: marriage, sex, salary, job status, industry. Based on these profile variables we clustered the customers into 16 clusters using K-means algorithm, see for example [10]. We performed two clusterings, one with all the customers and one with only those customers whose lifetime was larger than 1, as presented in III-1. In the latter clustering we assigned the trial customers to the nearest clusters to avoid information lost.

IV. RECOMMENDATION APPROACHES

We implemented several recommendation methods for presenting customers a selection of coupons from the set of available coupons.

A. Aggregate Level

1) Aggregate Popularity (P): As a reference and baseline, we aggregated coupon prints over all customers and dispenser during the aggregation period. For the aggregation period we used days= $\{3, 7, 14, 21\}$ number of previous days. After the aggregation we ranked the available coupons based on the popularity. The popularity based recommendations presents same coupons to all users at all dispensers at the same timestamp.

2) Dispenser Level (P_eq): For location based recommendation, we aggregated coupon prints over all customers during the aggregation period on a given dispenser. We used the same aggregation periods as in overall popularity, i.e. days= $\{3, 7, 14, 21\}$ number of previous days. The dispenser level popularity present same coupon sets to all users in a given dispenser at the same timestamp.

3) Customer's Previous Prints (P_f): We implemented a customer level filtering method for aggregate coupon prints. From the overall aggregate popularity we filtered away those coupons which had already been printed by the customer.

B. Segmentation Level

1) Customer Segments (U and U_o): We assumed that customers have heterogeneous preferences of different type of coupons. We analyzed the heterogeneity by simply grouping the customers into clusters. On this study, we tried cluster numbers from 6 to 20, and from sum of error distance found that 16 clusters are suitable for us. Two clustering were conducted, first based on demographic variables and whole customer base (U) and second (U_o) based on the demographic variables but the trial users were filtered away.

At customer segment level we aggregated coupon prints per coupon within customer segments and calculated the coupon popularities within segments. A customer will be presented top-n coupons in the segment she belongs to.

C. Individual Level

1) Customer Interest Based (I and I_U): As presented in III-2 for each customer we formulated a binary interest vector. With these interest vectors we aggregated interest vectors for each coupon in our coupon set from the training period. The coupon interest vector was calculated as an average over the interest vectors of the customers who printed the coupon (I). In addition, we also updated the customers interest vectors based on the coupon interest vector which were printed by the customer (I_U).

A customer is presented the top-n coupons from the available coupon set in a given timestamp based on the similarities of the coupon interest vectors and customer's interest vector.

2) Item-based Collaborative Filtering (item CF): We implemented a standard item-based collaborative filtering algorithm as presented in [11]. We calculated the item-similarity matrix in each day from the training period. We tried several training periods, days= $\{15, 25, 50, 200\}$. A customer is presented the top-n coupons from the available set in a given timestamp based on the cosine similarity metric.

V. Results

A. Evaluation Data and Metric

As for evaluating the recommendations presented in previous chapter the transaction records from period 1.2.-21.2.2010 were used of a test set. For each dispenser and timestamp a set of available coupons were calculated. On average 102.4 coupons were available on a given dispenser on a test set timestamp with standard deviation of 16.1 coupons. As mentioned above we updated our models once a day, the training data for each instant of time included all the transaction records from previous days, but not the previous records on the given day.

For measuring goodness of coupon recommendations we formulated a set of top-12 coupons for each transaction record timestamp, dispenser ID and customer ID information. [8] If the correct coupon belongs to the presented top-12 list we count it as a good coupon recommendation.

B. Coupon Recommendations

Table II summarizes the results for the discount coupon recommendations. Because we tried several training and aggregation periods, in Table II only one of these combinations is chosen. 21 days for the aggregation and 200 for the item similarity training. The overall popularity results reflect the results of dispenser level popularity and customer level popularity in different aggregation periods.

Table II
RATIOS OF THE CORRECT COUPON FOUND IN THE PREDICTED TOP-12. DAILY AVERAGE AND STANDARD DEVIATION ARE CALCULATED FROM DAILY RATIOS OF THE TEST PERIOD 1.-21.2.2010. POPULARITIES ARE CALCULATED FROM THE LAST 21 DAYS. INTEREST SIMILARITIES AND ITEM-BASED CF ARE CALCULATED FROM THE TRAINING SET FROM 200 PREVIOUS DAYS.

method	overall	daily average	daily std.
popularity	48.9%	48.6%	4.0%
popularity_eq	54.6%	54.7%	3.6%
popularity_f	35.7%	35.2%	4.0%
user_segment	36.4%	36.7%	3.6%
user_segment_o	36.6%	36.9%	3.5%
interest sim	10.9%	10.7%	1.9%
interest u	24.6%	24.7%	3.1%
item CF	13.0%	13.0%	1.6%

[8]From the dispenser implementation we cannot say which button location is best. Because each dispenser has at least 12 buttons, we chose top-12 to assimilate the idea of coupon appearing upfront.

The Figures 3 and 4 present the impact of different aggregation and training periods. In Figure 3 the overall popularity is compared with customer based segmentation popularity on different aggregation period. In Figure 4 the item-based CF and interest vector based recommendations are compared. The results show that customer based segmentation produces nearly identical results, the filtering of trial customers does not have effect. The Figure 3 shows that overall popularity produces weaker results when aggregation period is greater. However, the Figure 4 shows that both interest based recommendations and CF produces better recommendations when more training data is utilized.

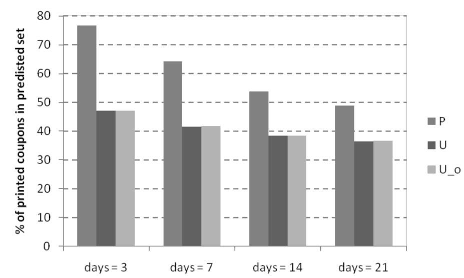

Figure 3. Ratios of the correct coupon found in the predicted top-12 during the whole test period. The different days refer to the period where the coupon popularities are calculated. P: most popular coupons without filtering, U: most popular coupons in customer segment when trial customers are filtered and U_o: most popular coupons in customer segment without filtering.

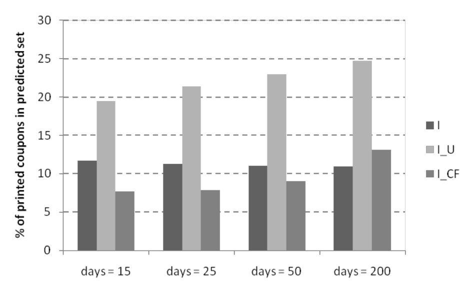

Figure 4. Ratios of the correct coupon found in the predicted top-12 during the whole test period. The different days refer to the period where interest and item similarities are calculated. I:Interest based similarity, I_U: Interest based similarity with an update and I_CF: item-based collaborative filtering.

Figure 5 summarizes the results from different days in the test period. The overall popularity, customer based segmentation popularity and interest similarity recommendations are compared. The figure 5 indicates that the distribution of coupon prints is very similar on daily basis and each recommendation approach is able to detect nearly constant proportion of the printed coupons.

VI. Conclusions

This paper has focused on personalization of discount coupons in coupon dispensers. The aim is to cut down the

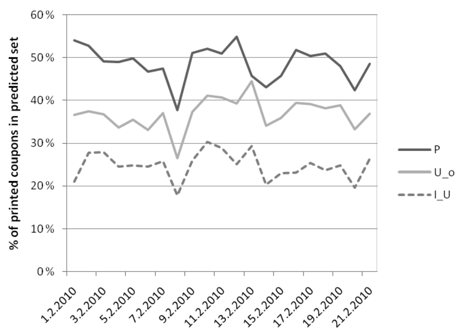

Figure 5. Ratios of the correct coupon found in the predicted top-12 on each day during the test period 1.-21.2.2010. P: most popular coupons without filtering, U_o: most popular coupons in customer segment without filtering, and I_U: most similar coupons based on customer similarities with an update.

customer's usage time in front of the dispenser. This work has not tackled nor measured the actual time customer's spent using the dispenser. Instead we have presented recommendation methods for personalizing the coupons upfront in the dispenser.

The result show that overall popularity is a good starting point for recommendation analysis. The aggregation of coupon prints from previous couple days on dispenser level produced the best results while methods without popularity based ranking did produce poorer results. The results may arise from the fact that no recommendation methods were actually implemented, and in the test set the coupons were printed by the customers because the coupons were already upfront and thus popular. However, the results also show that if more training data is available, the actual coupon prints reflect more customer interest than the interest the customers indicate themselves, the methods I_U versus I.

We admit we have not produced high quality results and maybe presented more questions than we have answered. However, this paper has shown that with basic methods and thoughtful background analysis the discount coupon recommendations are possible without content information. We believe that by presenting relevant coupons to the customers both the customer satisfaction increases and customer's usage time decreases.

A. Future research

Because the results were distorted towards the popularity of coupons, the future research could focus on presenting some less known coupons to the customers and evaluating the value of discovery to the customers. The actual implementation and empirical knowledge is needed.

Another research aspect is the possible extension from physical dispensers to the mobile solutions, which would implicitly give the provider information about the actual use of the coupons. The relevant feedback would also give valuable information about context of coupon prints and use. How will the weekday and dispenser location effect the customer's print behavior? How time depended the coupon types are, i.e. when to recommend coffee again? With digital coupons, also more specific use cases would be possible, for example shop owners aiming to find new customers could issue coupons that would be only valid for first time buyers.

References

[1] G. Shaffer and Z. J. Zhang, "Competitive coupon targeting," *Marketing Science*, vol. 14, no. 4, pp. pp. 395–416, 1995.

[2] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Modeling and User-Adapted Interaction*, vol. 12, pp. 331–370, 2002.

[3] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *Internet Computing, IEEE*, vol. 7, no. 1, pp. 76–80, 2003.

[4] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2008, pp. 426–434.

[5] Z. Abbassi, S. Amer-Yahia, L. V. Lakshmanan, S. Vassilvitskii, and C. Yu, "Getting recommender systems to think outside the box," in *RecSys '09: Proceedings of the third ACM conference on Recommender systems*. ACM, 2009, pp. 285–288.

[6] W. Chu and S.-T. Park, "Personalized recommendation on dynamic content using predictive bilinear models," in *WWW '09: Proceedings of the 18th international conference on World wide web*, 2009, pp. 691–700.

[7] J. Liu, P. Dolan, and E. R. Pedersen, "Personalized news recommendation based on click behavior," in *IUI '10: Proceeding of the 14th international conference on Intelligent user interfaces*. New York, NY, USA: ACM, 2010, pp. 31–40.

[8] P. E. Rossi, R. E. McCulloch, and G. M. Allenby, "The value of purchase history data in target marketing," *Marketing Science*, vol. 15, no. 4, pp. 321–340, 1996.

[9] U. Ben-Zion, A. Hibshoosh, and U. Spiegel, "The optimal face value of a discount coupon," *Journal of Economics and Business*, vol. 51, no. 2, pp. 159 – 174, 1999.

[10] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

[11] M. Deshpande and G. Karypis, "Item-based top-n recommendation algorithms," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 143–177, 2004.