

Products Recommendation for Mobile Devices

Ruyther Parente da Costa^{1*}, Caíque de Paula Pereira¹, Edna Dias Canedo¹

¹ Universidade de Brasília (UnB), BRAZIL

*Corresponding Author: ruyther@me.com

Citation: da Costa, R.P., Pereira, C.d.P., and Canedo, E.D. (2017). Products Recommendation for Mobile Devices. *Journal of Information Systems Engineering & Management*, 2(3), 16. doi: [10.20897/jisem.201716](https://doi.org/10.20897/jisem.201716)

Published: August 2, 2017

ABSTRACT

The mobile application market and e-commerce sales have grown steadily, along with the growth of studies and product recommendation solutions implemented in e-commerce systems. In this context, this paper proposes a recommendation algorithm for mobile devices based on the COREL (CustOmer purchase pREdiction model) framework. The proposed recommendation algorithm is a customization of the COREL framework, based on the complexity of the implementation associated with iOS mobile applications. Therefore, this work aims to customize a gift recommendation algorithm in the context of mobile devices using as main input the user preferences for the gifts recommendation in the Giftr application.

Keywords: recommendation algorithms, e-commerce, mobile software development, iOS, Giftr

INTRODUCTION

The electronic market in recent years has seen an intense growth in sales in Brazil, with the festive dates of the year (Mother's Day, Valentine's Day, Children's Day, Christmas Day, etc.) the time of year when demand by this type of business intensifies (Mendes, 2016). This scenario provides a business opportunity for software applications that offer the consumer the opportunity to gift someone on those festive dates by cross-referencing user profile data to be presented to offer the best related gifts.

To improve the user experience and increase their sales, online sales companies use product recommendation algorithms according to the characteristics of their consumers. There are several types of algorithms, and despite the success of some of them, most of it has problems. Therefore, it is important companies have a consistent and relevant algorithm to recommend products to their users (Gama, 2011).

One of the problems encountered in creating recommendation algorithms is that initially the system has little user information. This makes learning and algorithm performance difficult. Thus, it is necessary to use mechanisms that reduce the learning time of the algorithms and predict products the user might need based on the little available information (Gama, 2011).

The present work aims to recommend the gifts that best suit the user, based on information of his profile and personal tastes that he has in the Giftr application (Ruyther et al., 2017).

THEORETICAL BASIS

The recommendation systems are intended to assist in the suggestion of items, products, services and/or content, in a partial or fully automatic manner, according to the user's interest (Burke, 2002). These systems can provide information that will assist the user in making decisions about which items to choose, such as books,

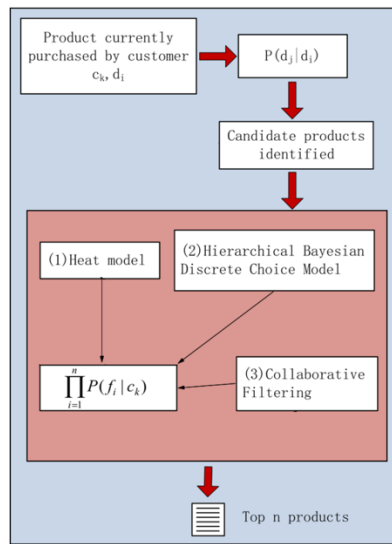


Figure 1. Workflow of COREL

music, movies, products, as well as referral systems can, based on a user’s profile, suggest items to the user directly without the intermediation of the same, such as gifts that the user may have interest based on his profile.

According to Resnick and Varian (1997), a recommendation system has as input data which the user grants, the system then uses them to make the recommendations and then directs them to the relevant recipients. The term “recommendation systems” came to replace the term “collaborative filtering”, after which collaborative filtering refers to a specific recommendation algorithm. In general, referral systems are referred to as systems that recommend a list of user products or systems that help users evaluate products.

RELATED WORK

Several studies in the field of e-commerce recommendation algorithms have been developed, and this in particular proposes a framework that aims to predict user behavior in the context of e-commerce (Qiu et al., 2015; Vasquez et al., 2016). The great differential presented in this article is that it aims to predict consumer behavior, in other words, the consumer’s preferences to buy some product in an ecommerce system. The article points out that through traditional algorithms there is no satisfactory execution of predictive tasks, so the article proposes a framework, COREL (CustOmer purchase pREdiction model), a solution capable of solving this very common challenge in the traditional business context.

COREL, in a more detailed way, has an initial step in identifying the products purchased by the consumer in the e-commerce system, after which this data will later be an important input to the subsequent steps that will try to predict, for example, what are the characteristics (Example: price, color, brand, etc.) of the product that the consumer believes is decisive at the moment of its acquisition.

$$P(d_j, d_i) = \frac{|d_i \cap d_j|}{d_i} \tag{1}$$

$$P(d_j, d_i) = \frac{|Thr(d_i) \cap Thr(d_j)|}{Thr(d_i)} \tag{2}$$

The second part of the initial stage is defined by the calculation of Equation 2, based on the likelihood model Equation 1 (Ponte and Croft, 1998), which represents an association in which the products d_i and d_j , that is, given the purchase of product d_i would be calculated the probability of the consumer to acquire the product d_j , being in Equation 1 the number of products purchased and $|d_i \cap d_j|$ the frequency that products d_i and d_j co-occur in the same e-commerce system. However, the author’s proposal to construct an association between categories in order to obtain a relation with the selected candidate products coming from several categories with a particular product, for this reason the author presents the Equation 1, where $Thr(d_i)$ represents the third-level categories of product d_i . At the end of this initial stage the identification of the candidate products is carried out, based on the execution of the previous stages.

In COREL’s main step, consumer preferences, already identified, are used to identify which candidate products are most likely to be purchased. For further refinement of the selected candidate products, the author investigates three product categories:

- Heat Model: it aims to predict the characteristics that the consumer may be interested in by calculating the popularity of these characteristics, based on the SVR (Support Vector Regression) model, using some characteristics as basis for the model, they are: Q_r , Q_s , Q_a And Q_u ;
- A Hierarchical Bayesian discrete choice model: the author proposes to develop a discrete choice Bayesian hierarchical model to calculate the probability of d_j choosing the d_j product based on its brand preference and price sensitivity;
- Collaborative Filtering: the model aims to predict the evaluation that the consumer would give to some product, based on the other consumers of the system that has similar tastes to his.

METHODOLOGY

A systematic mapping was performed with the objective of identifying articles related to the recommendation algorithms based on the user profile found in the literature that answer the research questions raised in the application of this technique.

Research Questions

The systematic review attempted to answer the follow research questions:

- **Q1:** What is recommendation algorithm?
- **Q2:** Which are the recommendation algorithms in the literature?
- **Q3:** What are the most commonly used recommendation algorithms in the literature for selecting a desired item?
- **Q4:** The existing algorithms can be used or must customize it and create a selection algorithm?

Research String

To answer the research questions from the previous section, the definition of the string is a necessary step to perform the search on some digital scientific bases. The construction of the string is based on the guidelines defined by Keele (2007), which consists of identifying keywords from search queries and uses AND connectors to combine keywords and ORs to match synonyms.

The following keywords were considered to search the selected databases in Portuguese:

- Algoritmo;
- Recomendação;
- Algoritmo de Recomendação;
- Algoritmo de Recomendação baseado no perfil;
- E-commerce.

And the following keywords in English were considered:

- *Algorithm*;
- *Recommendation Algorithm*;
- *Recommendation Algorithm based on profile*;
- *E-commerce*

The search string created for the search was as follows:

(algorithm OR algoritmo) AND (recommendation based on profile OR recomendação baseado no perfil OR recommendation OR recomendação) AND (e-commerce)

Results

According to the Systematic Review process carried out, the research questions were answered:

- **Q1:** recommendation systems are widely used in many applications to suggest services and information to potential consumers. Services like Amazon.com, Netflix.com, Half.com have used referral systems and increased the loyalty of their consumers. Good recommendation algorithms are necessary for the operation of these recommendation systems (Huang Daniel Zeng, 2004).
- **Q2:** the recommendation algorithms that were found are content-based, collaborative recommendation, hybrid approach, user-based algorithm, item-based algorithm, and dimensionality-reduction algorithm.
- **Q3:** algorithms based on user behavior.
- **Q4:** the existing algorithms are not enough to fit the context of this work, however some can be availed. Therefore, it is necessary to customize the recommendation algorithm to the desired context.

DEVELOPMENT

The Giftr application was created in the BEPiD (Brazilian Education Program for iOS Development) project (Universidade Católica de Brasília, 2016) with the idea of helping people give gifts to each other. The solution found by the team was to develop a social network where each user registers their favorite products, tastes and sizes (shoes, blouses, etc.), and with this data the user has the possibility to give another through the application.

The main functionalities of the application already developed in BEPiD were: search of products and users, sending of notifications and control of data of the user profile.

The functionality of the search application, both user and product has a fundamental role in the application, because through them users can find other users and invite them to be your friends. Already the search for products allows the user to find products in general, based on the products available from the Lomadee API (Lomadee, 2016a), enabling the user to purchase products and evaluate the products, ranging from zero to five points, to show in the system how much the user wants to be presented with that product.

The data control functionality of the profile allows the user to change and add personal information of the user, this being the means that the same has to register their tastes, fundamental for the operation of the algorithm of recommendation, and the measures, the size of footwear used by him. The registration of the tastes occurs through the entry by the user of a string that represents a taste of yours, for example, “iPhone”, and later inform which category of the search is associated with preference, for example “Cellular and Smartphone”.

Lomadee Plataforma

Buscapé (Company, 2016) provides some very robust platforms, among which is the Lomadee (Lomadee, 2016b), which provides several APIs for data access available in the Buscapé system.

Lomadee provides several APIs (Lomadee, 2016b), they are offer API, coupon API, Report API.

The API used in the Giftr application is the one of the Offers of the Lomadee platform, because through it there is the possibility to retrieve data from categories, products, offers and evaluations of users and stores of Buscapé, fundamental to the Giftr application and to the recommendation algorithm of gifts. This API provides several types of query for data recovery and among them the ones that will be used in this work are Find Category List, Top Products, Find Products List, View User Ratings, Top Offers, Find Offer List.

Mobile Devices Gift Recommendation Algorithm

The algorithm proposed in this article will be based in another article, the Predicting Customer Purchase Behavior in the E-commerce Context (Qiu et al., 2015), which will be customized to be in accordance with the mobile context, more precisely the application for iOS Giftr.

The framework COREL was proposed for the e-commerce context, which aims predict the customer behavior and recommend products based on that prediction. The context proposed for this algorithm is a mobile application that it helps people give a present to the other, recommending products based on the user profile. The scope presented in the article is much larger than the one proposed for this article, in a context also different, and for these reasons, COREL will be customized to suit the context of a mobile application, Giftr.

Based on the COREL flow, the algorithm proposed by this work will follow a similar flow with a hybrid recommendation systems approach. Figure 2 exemplifies the flow of the proposed algorithm, and the other subsections detail what happens in each step.

1. Categorize the Products Rated by the User

The first step of COREL, the “product currently purchased by customer c_k , d_i ”, which consists of the verification of the product d_i purchased by the consumer, allowing the framework a basic product for the probability calculations that will be performed later.

This context, however, differs greatly from the one presented in this paper, after all the Giftr application aims to help people to present others with a social network, and does not fit into the category of an e-commerce system in which COREL operates.

The proposed algorithm, instead of verifying the product (d_i) that the consumer (c_k) has acquired, identifies the product (p_i) that the user evaluated in the application (ranging from zero to five). This evaluated product (p_i) is then written in a list of products evaluated by the user (l_p) for later use in the algorithm.

The Lomadee API returns several product attributes, and they are general attributes for any platform product. Thus, the choice of which attribute of the product will be stored is important and determinant for the proposed algorithm, so the attributes selected were: product name (pn), product category (pc), minimum price ($Qpmn$), maximum price ($Qpmx$), average score of product ratings (Qs) and total number of comments (Qr).

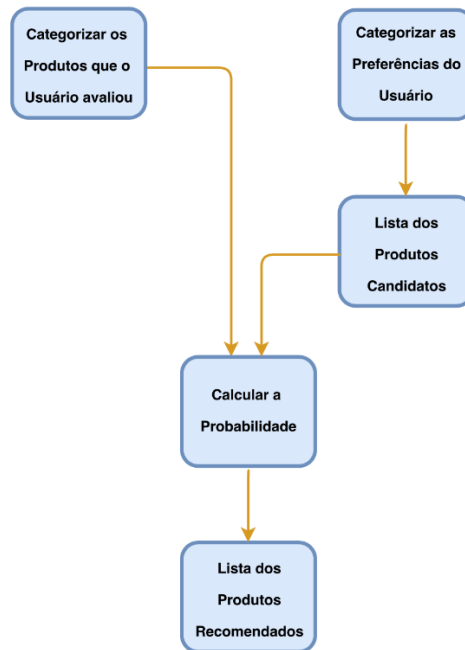


Figure 2. Workflow of the proposed algorithm

2. Categorize the User's Preferences

In COREL, user preferences are identified in a predictive way, that is, through iterative steps (1) Heat Model, (2) A hierarchical Bayesian Discrete Choice Model and (3) Collaborative Filtering. The model seeks to predict the consumer's tastes for a given product from product data that the consumer has already acquired, product preference data (rating numbers, average ratings, etc.) reported by the consumer which it believes to be of greater relevance and of consumers who have similar tastes, to predict the preferences that a certain user of the system will have when purchasing products.

The context of the previous paragraph presented is not the same as the one found in Giftr, after all the user will already inform their tastes based on products of the most varied categories, such as cellular, computer, among others. These preferences will be used to perform this step of the recommendation algorithm, and no tasting prediction method is required by COREL.

The user should enter a string (ps) that represents his preference and the category (pc) associated, according to a category of products of the buscapé. Thus, these data will serve as inputs to the next step of the proposed recommendation algorithm, and for this reason will be stored for later use.

3. Products Candidates List

In this step will be done the product listing (p_j) that will be used for the probability calculations of the next step. For the listing of the candidate products it is necessary to inform two important data for the use of the Lomadee Offers API, using the Find Products List query type, the keyword (ps) and the category (pc) of the preference informed by the user in the application, so the API returns the existing products in the Buscapé associated with ps and pc .

In the Lomadee API, the data is entered through a url, and the products that are returned by the API do not have a defined quantity, so it is necessary to define the quantity of products that the list of candidate products will have, taking into account since the algorithm will run in a mobile application with a hardware limitation that can impact the other steps of the algorithm.

The user of the Giftr application has the possibility to inform several preferences and several categories associated with them, therefore, the scope of the product listing is limited to a category ($pc = 1$), however, having the possibility to have one or more keywords associated with this category ($s \in Z | s > 1$), for example, "iPhone" and "Samsung Galaxy" as keywords in the "mobile and smartphone category."

The " n " quantity of products will be defined in the next steps in an empirical way, mainly taking into account the performance and usability for a better user experience, since this decision is totally tied to the implementation of the algorithm in the application.

At the end of this step are listed " n " candidate products (w).

4. Probability Calculation

This step consists in briefly calculating the probability of the user (c_k) becoming interested in the product (p_j), comparing the characteristics $Qpmn$ and $Qpmx$ of this product with that of the product evaluated by the user (p_i), both of the same category (pc).

The calculation of the probability proposed by this article differs greatly from the base article of this proposal, since in it the calculation is performed using the methodology represented in [Figure 1](#) by calculating the probability $P(d_j|c_k)$. In the calculations of the algorithm proposed in this work, there will be no specific equation for the calculation of this probability, in the case it will be a sequence of steps that will define the products with the highest probability that the user will be interested.

To do so, this step will be subdivided into three substeps so that the products with the highest probability are listed, they are:

a. Make a Price Comparison Between the Favored Product and the Candidate

The first step in this substep is the search of the stored data of the user's favorite product (p_i), output of the step, as they will be the basis for comparisons with the candidate products, as well as to retrieve the data of all products from the product list candidates (w).

The parameters that will be used are rescued from the product evaluated by the user: p_n , $Qpmn$ and $Qpmx$. The first parameter will be necessary for the product identification, the second and third are the parameters that best show the characteristic of this product, among other parameters that the API returns in the query and the parameters that will be used for pc will be the same as p_n .

Given the parameters that will be used, the calculations that will be carried out for the comparison of p_j and p_i are:

$$PQmn = \frac{Qpmn(p_j)}{Qpmn(p_i)} \quad (3)$$

where $Qpmn(p_j)$ is the minimum price of p_j , $Qpmn(p_i)$ is the minimum price of p_i , $PQmn$ the ratio of $Qpmn(p_j)$ to $Qpmn(p_i)$, and

$$PQmx = \frac{Qpmx(p_j)}{Qpmx(p_i)} \quad (4)$$

where $Qpmx(p_j)$ is the maximum price value of p_j , $Qpmx(p_i)$ is the value of the maximum price of p_i , $PQmx$ is the ratio of $Qpmx(p_j)$ to $Qpmx(p_i)$.

A separate calculation for the $\min(PQmn)$ and $\max(PQmx)$ prices will be carried out, where the calculations are intended to show how much the price, p_i and p_j , represent in relation to p_i , and thus identify the p_j that similar to p_i prices.

This sub-step is then terminated and the generated data is stored for use in the next substeps of the algorithm, so the list of candidate products (w) is updated with the values of the proportions calculated from Equations 3 and 4.

b. Filter the Candidate Products that Deviate from the Minimum and Maximum Price of the Favorite Product

This sub-step has as inputs the updated list w , with the values of the proportions of the minimum and maximum prices of the products p_j . Thus, the objective of this substep is to filter the products that disagree "x" per cent from the base prices (minimum and maximum) of p_i and exclude those that exceed a limit percentage value.

To calculate the proportional percentage that the parameters $Qpmx$ and $Qpmn$ of p_j have in relation to the same parameters of p_i , it is necessary to perform the following calculations:

$$PcQmx = PQmx \times 100 \quad (5)$$

$$PcQmn = PQmn \times 100 \quad (6)$$

Equations 5 and 6 indicate the proportional percentage of the minimum and maximum price of p_j in relation to p_i . For filtering of the products p_j to occur, a threshold value (pl) is required both downward and upward of the base value, for example, ten percent up and down one hundred percent of the minimum and maximum price of p_i , and the definition of the value of pl depends on actual tests with the algorithm in Giftr, since it can impact performance if it is too large and if it is too small it may not return any product. For these reasons, the threshold percentage value pl will be defined in the next steps.

Then with the percentage values $PcQmx$ and $PcQmn$ for each product, the classification of those that comply with the percent limit pl . If any product has $PcQmx$ and $PcQmn$ outside the percent limit pl , it is excluded from the list of candidate products, if only one of the percentage values is not in the limit pl the product is not excluded from the list w , as in the case that $PcQmx$ and $PcQmn$ are within the limit, as shown in [Table 1](#).

Table 1. Candidate product exclusion truth table

$PcQmx$	$PcQmn$	$PcQmx \vee PcQmn$	Result
1	1	1	It's not excluded
1	0	1	It's not excluded
0	1	1	It's not excluded
0	0	0	It's excluded

If there are many candidate products at the end of this filtering, it may be necessary to define a threshold number so that there will be no performance problems in the algorithm, which will be defined in the future because several tests are required for such a definition.

c. Elaborate the Rank of the Candidate Products Based on the Minimum and Maximum Prices

Based on the calculations of the previous substage of the price ratio $PcQmx$ and $PcQmn$ of p_j with respect to p_i , the list w contains the all of the p_j in disorder. The purpose of this substage is to sort the list based on the $PcQmn$ and $PcQmx$ data.

$$Pcm = \frac{PcQmx + PcQmn}{2} \quad (7)$$

Since there are two distinct data, $PcQmn$ and $PcQmx$, for ordering the list in a way that is more optimized, the arithmetic mean of these two Pcm values will be given so that only one value is present for the comparison at the moment of descending ordering of the products, as shown in Equation 7.

5. List of Recommended Products

The purpose of this step is to reorder the list w based on the comparison of two more parameters of p_j and p_i , Qs and Qr . The motivation of this reordering is to give more credibility to the ordering of the products in the list, based on the data that Lomadee makes available in its API.

The parameter Qr informs the amount of comments that a product obtained in Buscapé, and can be used as a way to give credibility to the value given by Qs , that is if one product has Qs equal to 9.0 and another also has 9.0, a higher value of comments (Qr) will have a greater relevance in relation to the other.

$$Pd = \frac{Qs}{Qr} \quad (8)$$

The parameter Pd then indicates the credibility to the value of Qs , in the case the closer to zero Pd is, the greater the credibility of Qs , since Qr tends to be a larger value. Then for the reordering the Pd will be used, with the list w already found, so that the reordering will be re-done without taking into account the one performed by the step of the previous algorithm, the value of Pd will be added to Pcm :

$$Pw = Pd + Pcm. \quad (9)$$

Pw is the base value for the decreasing reordering of list w , which takes into consideration Pcm from the first ordering of the third substep. At the end of this step, the list w has the products p_j in descending order of importance to be recommended to the user.

CONCLUSION AND FUTURE WORK

Because this application does not use the data received to process this work, it was possible to perceive the relevance of a product recommendation algorithm in the context of e-commerce applications. Therefore, the solution of this problem through an algorithm has been intensely discussed in scholarly works that seek to create better algorithms as presented in the Systematic Review.

Through the Systematic Review, possible gift recommendation solutions that take into account the user profile were investigated. Among the solutions found, the one that best matches the context of this work is the COREL framework.

The Giftr application is an iOS platform application that uses the Lomadee API, provided by Buscapé, where you can search for products, subscribe to user preferences and with that input to recommend the best products to the user.

This article proposes the customization of the COREL framework, in the context of mobile applications in which Giftr is found, using as main input the user preferences. This work has not yet been completed, more activities will be held to complement it.

REFERENCES

- Burke, R. (2002). *Hybrid Recommender Systems: Survey and Experiments*. Fullerton. Company, B. (2016). [Online]. Available in: <http://developer.buscape.com.br/portal>. [Accessed: 26- Nov- 2016].
- Gama, R. et al. (2011). *Algoritmo de Recomendação Baseado em Passeios Aleatórios num Grafo Bipartido*.
- Huang Daniel Zeng, H.C.Z.A. (2004). *Comparative Study of Recommendation Algorithms in E- Commerce Applications*.
- Keele, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*.
- Lomadee, A.d.O. (2016a). [Online]. Available in: <http://developer.buscape.com.br/portal/lomadee/api-de-ofertas/recursos>. [Accessed: 6- Dec- 2016.]
- Lomadee, P. (2016b). [Online]. Available in: <http://developer.buscape.com.br/portal/lomadee>. [Accessed: 26- Nov- 2016].
- Mendes, R. (2016). *Os números do mercado de e-commerce*. [Online]. Available in: <http://www.profissionaldeecommerce.com.br/e-bit-numeros-do-e-commerce-no-brasil/> [Accessed: 26- Nov- 2016].
- Ponte, J.M. and Croft, W.B. (1998). *A Language Modeling Approach to Information Retrieval*.
- Qiu, J., Lin, Z., Li, Y. (2015). *Predicting customer purchase behavior in the e-commerce context*. Springer Science. [S.l.].
- Resnick, P. and Varian, H.R. (1997). *Recommender Systems*.
- Ruyther, P.C., Pereira, C.d.P. and Dias Canedo, E. (2017). *Recomendação de Presentes em Dispositivos Móveis*. In: *12ª Conferência Ibérica de Sistemas e Tecnologias de Informação (CISTI)*.
- Universidade Católica de Brasília. (2016). Distrito Federal DF, Brasil, *Brazilian Education Programm for iOS Development*.
- Vasquez, M.Y., Sexto, C.F., Rocha, A. and Aguilera, A. (2016). Mobile phones and psychosocial therapies with vulnerable people: A first state of the art. *Journal of medical systems*, 40(6).