Mobile Gift Recommendation Framework
A COREL Framework Approach

Caíque de Paula Pereira[1], Ruyther Parente da Costa[2] and Edna Dias Canedo[2]
[1]*Faculty of Gama (FGA), University of Brasília (UnB), Brasília-DF, Área Especial de Indústria, Projeção A – P.O. Box 8114 – CEP 72.444-240, Brazil*
[2]*Department of Computer Science – Edifício CIC/EST – Campus Darcy Ribeiro, Asa Norte - University of Brasília (UnB),*

Keywords: E-commerce, Recommendation Algorithms, Mobile Software Development, iOS, Giftr.

Abstract: This paper proposes a recommendation algorithm for mobile devices based on the COREL framework. In this context, the mobile application market and m-commerce sales have grown steadily, along with the growth of studies and product recommendation solutions implemented in e-commerce systems. 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. This algorithm has been tested through three cycles of tests and improved during it, the results suggest that the algorithm presents a good performance and gifts results based on the user preferences.

1 INTRODUCTION

In order to improve the user experience and increase their sales, e-commerce companies use product recommendation algorithms according to the characteristics of their consumers. There are many types of algorithms, and despite the success of some of them, most of them have problems. Therefore, it is important that companies have a consistent and relevant algorithm to recommend products to their costumers (Gama et al., 2011).

In recent years, the online sales have seen an intense growth in sales in Brazil and the rest of the World. Festive dates of the year like Mother's Day, Valentine's Day, Children's Day, Christmas Day and others are the time of year in which the demand for this type of trade intensifies e-commerce(Mendes, 2016). This scenario has provided 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 offer the best related gifts.

Initially, one of the problems encountered in creating recommendation algorithms is that, the system has not much user information. That hinders the learning and performance of the algorithms. Thus, mechanisms that reduce the learning time of the algorithms and prediction based on the little available information are necessary (Gama et al., 2011).

In the 90s, the e-commerce began when the first sales websites were created. Initially, the volume of transactions was very low. But the change in the world market made it become the largest and most voluminous way to market products (do Nascimento et al., 2009). According to a survey released by Buscapé in 2015, billions of Brazilian reais are collected on commemorative dates for e-commerce sales. The revenue for Christmas in 2015 was 7.40 billion Brazilian reais through purchases made on the Internet(de Souza, 2013).

Studies have been done in the area of e-commerce, especially about algorithms of recommendations, widely used in web systems. A good example of a study that proposes an improvement of recommendation systems using techniques that aim to predict the behavior of the user is the article Predicting Customer Purchase Behavior in the E-Commerce Context, later to use it as input for the recommendation of Products itself (Qiu et al., 2015).

The Giftr application is the study object of an initiative in this context. This application was designed and developed by Ruyther Costa, Caíque Pereira, Caio Sanchez and Victor Bruno, during the BEPiD project in the period of February 2 and December 11,

de Paula Pereira, C., Parente da Costa, R. and Dias Canedo, E.
Mobile Gift Recommendation Framework.
DOI: 10.5220/0006792806570663
In *Proceedings of the 20th International Conference on Enterprise Information Systems (ICEIS 2018)*, pages 657-663
ISBN: 978-989-758-298-1

2015.

Based on information of the user profile and personal preferences in the Giftr application, this research seeks to recommend gifts that best suit the user. The Buscapé engine has been chosen because of its popularity and available API in the context of brazilian e-commerce, which will support the Giftr application and the gift recommendation algorithm.

A common issue found in e-commerce stores and sales applications is the creation of good products by means of recommendation algorithms for their users. These algorithms help both the user experience and the increased sales of the companies.

This work proposes the following question due to this difficulty:

Q1. How to create a gift recommendation algorithm that fits the demand of a mobile App?

This work aims to contribute improving recommendation algorithms answering this question. The data collected through it can be used for future work related to the algorithms of recommendations focused on the e-commerce for gifts. This work general objective is to develop a gift suggestion algorithm that recommends the best products to the user based on their profile.

Investigating possible gift recommendation solutions that consider the user profile, substantiating the adopted solution with other algorithms solutions are the secondary objectives of this work.

1.1 Related Work

Some remarkable studies have been developed in the field of e-commerce recommendation algorithms, and this in particular proposes a framework that aims to predict user behavior in the context of e-commerce (Qiu et al., 2015). The big idea 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 e-commerce system. The article points out that through traditional algorithms there is no satisfactory execution of predictive tasks, so the article proposes a framework, COREL, a solution capable of solving this very common challenge in the traditional business context.

COREL, the framework proposed in this study is consists of two stages. The first stage is making an association between products by raising what is common among them and from these data to predict the motivations that lead the consumer to buy a particular product, and then build a list of products candidates for purchase by this consumer. The second stage is to predict the main characteristics that the consumer will be interested in a particular type of product and through these data define the products in which the consumer will be interested, based on the list of candidate products generated at the end of the first stage.

2 DEVELOPMENT

The Giftr application was created in the BEPiD (de Braslia, 2016) project 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, t-shirts, etc), and with this data the user has the possibility to give another through the application.

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 thus invite them to be your friends. The search for products allows the user to find products in general, based on the products available from the API of the Lomadee (Lomadee, 2016b), enabling the user to make the purchase of products and evaluate the products, with a variation of 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 Buscapé is associated with preference, for example ”cellular and smartphone”.

2.1 Lomadee Platform

Buscapé (Company, 2016) offers some very robust platforms, among which is Lomadee (Lomadee, 2016c), which provides several APIs for data access available in the Buscapé system. Lomadee offers several APIs (Lomadee, 2016a), they are:

- Offers API: it allows to retrieve data of categories, products, offers and evaluations of users and stores of Buscapé;
- Coupon API: enables you to query for active coupons on the Lomadee platform;
- Reporting API: Enables the retrieval of transaction or commission data in detail.

The API used in the Giftr application is that of the Offers on the Lomadee platform, because through it there is the possibility to retrieve data from categories, products, offers and evaluations of Buscapé users and stores, which are fundamental to the Giftr application and for the recommendation algorithm of gifts operation presented in this paper. This API provides several types of query for data recovery and among them the ones that are used are:

- Find Category List: returns detailed information of existing product categories in Buscapé and Lomadee;
- Top Products: returns the best products from Buscapé and Lomadee, processed and filtered by a exclusive technology of the platform;
- Find Products List: lists with detailed product information on Buscapé and Lomadee;
- View User Ratings: returns general user rating data about a specific product;
- Top Offers: returns the most searched products in Buscapé/Lomadee;
- Find Offer List: returns a list of the sites that are offering the product.

In addition to the outputs that each type of query returns, it is necessary to have a well-structured input so that the results are correct, the complete description of the inputs and outputs of each query available on the website of the Lomadee platform (Lomadee, 2016b).

2.2 Gift Recommendation Algorithm on Mobile Devices

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 accordance with mobile applications.

The framework COREL was proposed for a 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 figure 1 shows the flow from the proposed algorithm with a hybrid approach (Section **??**), little similar to the one that COREL uses, and the subsections below detail each step.

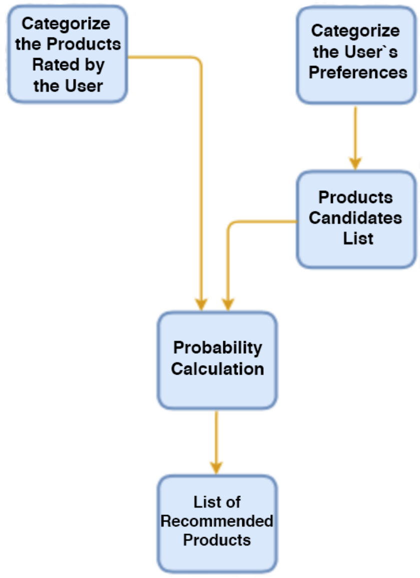

Figure 1: Proposed Algorithm with a hybrid approach.

2.2.1 Categorize the Products Rated by the User

In the first step from COREL, the "product currently purchase by customer c_k, d_i", which consists of verifying the product d_i bought by the purchased by the consumer c_k, given to the framework a base product that will allow the probability calculations to be performed later. This context, however, differs a lot of the one that the proposed algorithm is, because the main goal is to recommend gifts based on the user profile through a mobile application, the Giftr.

Having this in mind, the proposed algorithm, instead verify the product (d_i) that the consumer (c_k) purchased, identify the product (p_i) that the user rated in the mobile application, with a range from zero to five. The product rated (p_i) is wrote in a user‘s purchased products list (l_p) for further use in the algorithm.

The Lomadee API returns many product‘s attributes, usual from all products listed in the platform. Choose the right attributes to store is important and determinant as an input to the proposed algorithm, so the ones selected are: product name (p_n), product category (p_c), minimum price ($Qpmn$), maximum price ($Qpmx$), user average rating (Qs) and number of comments (Qr).

2.2.2 Categorize the User‘s Preferences

In COREL, the user‘s preferences are predict identified, in other words, through the interactive steps (1) Heat Model, (2) A hierarchical Bayesian Discrete Choice Model and (3) Collaborative Filtering, shown in the figure 1. The model seeks to predict the tastes that the consumer will have for a given product from data of products that the same has already acquired, of product preferences data (number of comments, user average rating, etc.) reported by the consumer that is believed to be of greater relevance and consumers who have similar tastes, to predict the preferences that certain user of the system will have at the moment of purchase of products.

The context of the previous paragraph is not the same that Giftr has, after all the user will inform his tastes based in products of different categories, as smartphone, computer, among others. This preferences will be use to make this step of the recommendation algorithm, which does not have any method to predict the tastes of the user as COREL.

As presented in section 2.2.1, the user must insert a string (p_s) which represents the preferences of the user and the category (p_c) associated, based on the categories from Buscapé (Buscapé, 2016). In this way, this data will serve as input to the next step of the recommendation algorithm, and for this motive it will be saved.

2.2.3 Products Candidates List

In this step will be fulfilled the products listening (p_j) that it will be used for the calculations in the next step. To list the candidates products is need to inform two important data for the use of the Offers API of Lomadee, using the consult API called "Find Product List", the keyword (p_s) and the category (p_c) of the preference informed by the user in the mobile application to the API returns the existent products in Buscapé associated with p_s and p_c.

In the Lomadee API, the data input is made through a url and the products that are returned using the API does not have a defined quantity, therefore it is necessary define in a empirical way the quantity of products that will be in the candidates products list, given that the algorithm will run in a mobile application with a limited hardware.

The Giftr user has the possibility to inform many of his preferences and categories associated it, for this reason, the scope to the products listing is limited a one single category ($p_c = 1$), nevertheless, having the possibility of have one or more keywords associated to this category ($s \in Z | s > 1$), for example, "iPhone" and "Samsung Galaxy" as keywords and the category being "cellphone and smartphone".

In the final of this step, there are "n" candidates products listed w.

2.2.4 Probability Calculation

This step consists, briefly, in the calculation of the user (c_k) to be interested by the product (p_j), comparing the characteristics $Qpmn$ and $Qpmx$ of this product with the one rated product (p_i), both in the same category.

The probability calculation proposed in this article differs a lot of the base article (Qiu et al., 2015), because in it the calculation is accomplished using a methodology shown in the figure 1 through the probability calculation of $P(d_j|c_k)$. In the calculations of the proposed algorithm in this article, it will not be a specific equation for the probability calculation, in this 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 sub steps so that the products with the highest probability are listed, they are:

1. Make a price comparison between the favored product and the candidate;
2. Filter the candidate products that deviate from the minimum and maximum price of the favored product;
3. Elaborate the rank of the candidate products based on the minimum and maximum prices.

2.2.5 Make a Price Comparison between the Favored Product and the Candidate Product

The first step in this sub step is the search of the stored data of the product favored by the user, output from the step described in the section 2.2.2, because they will be the basis for comparisons with the candidate products, as well as to retrieve the data of all products from the candidate product list (w), output of the sub step described in the section 2.2.3.

The parameters that will be used are from the product evaluated by the user: p_n, $Qpmn$ and $Qpmx$. The first parameter will be necessary for the identification of the product, 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 pn.

Given the parameters to be used, the calculations to be performed for the comparison of p_j and p_i are:

$$PQmn = \frac{Qpmn(p_j)}{Qpmn(p_i)} \tag{1}$$

where,

$Qpmn(p_j)$ is the minimum price of p_j, $Qpmn(p_i)$ is the minimum price of p_i, $PQmn$ the proportion of $Qpmn(p_j)$ in relation to $Qpmn(p_i)$, and

$$PQmx = \frac{Qpmx(p_j)}{Qpmx(p_i)} \tag{2}$$

where,

$Qpmx(p_j)$ is the maximum price of p_j, $Qpmx(p_i)$ is the maximum price of p_i, $PQmx$ the proportion of $Qpmx(p_j)$ in relation to $Qpmx(p_i)$.

A separate calculation for the minimum (*PQmn*) and maximum (*PQmx*) prices shall be performed, where the calculations are intended to show how proportionally the maximum and minimum price of p_j in relation to p_i, and thereby identify the p_j which most closely resemble the prices of p_i. This sub step is then terminated and the generated data is saved for use in the next substeps of the algorithm, in this way the list of candidate products (w) is updated with the values of the proportions calculated from equations 1 and 2.

2.2.6 Filter the Candidate Products that Diverge of the Minimum and Maximum Price of the Favored Product

This sub step has as input 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 sub-step is to filter products that diverge from "x" percent of the base (minimum and maximum) prices of p_i and exclude those that exceed a threshold percentage value.

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

$$PcQmx = PQmx \times 100 \tag{3}$$

$$PcQmn = PQmn \times 100 \tag{4}$$

The equations 3 and 4 indicate the proportional percentage of the minimum and maximum price of p_j in relation to p_i. In order for filtering of products p_j to occur, a threshold percentage value (p_l) is required both downward and upward of the base value, such as, for example, ten percent up and down of the one hundred percent of the minimum and maximum price of p_i, and the definition of the value of p_l is made empirically.

Then with the percentage values *PcQmx* and *PcQmn* for each product, the classification of those that comply with the percent limit p_l is carried out. If any product has *PcQmx* and *PcQmn* outside the percent limit p_l, it is excluded from the list of candidate products, if only one of the percentage values is not in the limit p_l the product Is not excluded from the w list, as is the case that *PcQmx* and *PcQmn* are within the limit, as shown in Table 1.

Table 1: Truth Table.

$Qpmn$	$Qpmx$	$Qpmn \bigvee Qpmx$	**Result**
1	1	1	It is not excluded
1	0	1	It is not excluded
0	1	1	It is not excluded
0	0	0	It is excluded

If there are many candidate products at the end of this filtering, it may be necessary to define a limit number so that there are no performance problems in the algorithm, this number must be defined empirically.

2.2.7 Elaborate the Rank of the Candidate Products based on the Minimum and Maximum Prices

Based on the calculations of the previous sub step of the price ratio $Qpmn$ and $Qpmx$ of p_j with respect to p_i, the list w contains the p_j all disordered. The purpose of this sub step is to sort the list based on the *PcQpmn* and *PcQpmx* data.

$$Pcm = \frac{PcQmx + PcQmn}{2} \tag{5}$$

Since there are two distinct data, *PcQmn* and *PcQmx*, in order to sort the list in a way that is more optimized, the arithmetic mean of these two values (*Pcm*) will be given so that only one value For the comparison at the time of the descending ordering of the products, as shown in the equation 5.

2.3 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 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 credence to the value given by Qs, that is, if a product has Qs equal to 9.0 and another one has 9.0, what has a higher value of comments (Qr) will have a greater relevance in relation to the other.

$$Pd = \frac{Qs}{Qr} \tag{6}$$

The parameter *Pd* then indicates the credibility to the value of *Qs*, in case the closer to zero *Pd* is, the greater the credibility of *Qs*, because *Qr* tends to be a larger value. Then for the reordering will be used *Pd*, plus the list *w* already found, so that the reordering is 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*:

$$Pw = pd + Pcm. \tag{7}$$

Pw is the base value for descending reordering of the *w* list, which takes into account *Pcm* of the first ordering of the third sub step described in Section 2.2.7. At the end of this step, the *w* list has the products p_j in the order of importance to be recommended to the user.

3 ALGORITHM VALIDATION AND VERIFICATION

The following results were found using the methodology and steps defined, respectively, (de Paula Pereira et al., 2017) and (da Costa et al., 2017).

The type of tests chosen to validate the algorithm were gray and black box tests. Each test case has been defined, refined, executed and documented. They were executed on the iPhone 5 simulator of Xcode using iOS 10.3. The Table 2 shows how the results are represented by color in tables 3 to 5.

Table 3: Results representation.

Color	Test Case Result
green	Test case passed
yellow	Test case passed but had a warning
red	Test case did not pass or No results

The Table 3 shows the results from the first cycle of tests. The first cycle contains tests of gray and black type. The objective of the fist cycle was to test different numbers of p_l, interests, rated products with the same or different categories. The values of p_l that were tested were 15, 20 and 25. And the quantities of interests and rated products were none, 1, 3 or 10.

After the execution of the first cycle, 55% of the test cases passed. Also, 40% did not pass or had no results and 5% passed, but had a warning.

In TC-11, a bug was found when you add interests with the same name more than one time and with different categories. In addition to this, in TC-16 there was an error on SQLite on the first algorithm execution. Also, with this cycle, it was possible to observe that when the categories are totally different, there are no results.

Other bugs not related to the algorithm itself were found and fixed after this first cycle of tests with the bugs mentioned above. An important conclusion from those test cases was that the number of results of the recommendation for each test were lower than expected. Most of the results were only one recommendation to the user.

Table 4: First cycle of tests.

TC-1	TC-2	TC-3	TC-4	TC-5	TC-6	TC-7	TC-8	TC-9	TC-10
red	red	green	green	green	green	green	green	green	red
TC-11	TC-12	TC-13	TC-14	TC-15	TC-16	TC-17	TC-18	TC-19	TC-20
green	red	green	green	yellow	red	red	red	red	green

The Table 4 shows the results from the second cycle of tests. The second cycle contains tests of gray and black type. The objective of the second cycle was to test higher numbers of p_l, increase the quantity of interests and rated products with the same or different categories. The values of p_l that were tested were 20, 25 and 30. And the quantities of interests and rated products were none, 1, 3, 10 or 30.

Table 5: Second cycle of tests.

TC-1	TC-2	TC-3	TC-4	TC-5	TC-6	TC-7	TC-8	TC-9	TC-10
green	yellow	green	green	red	green	green	red	green	green
TC-11	TC-12	TC-13	TC-14	TC-15	TC-16	TC-17	TC-18	TC-19	TC-20
green	green	green	green	green	red	green	green	green	green
TC-21	TC-22	TC-23	TC-24	TC-25	TC-26				
yellow	green	green	red	green	green				

After the execution of the second cycle, 76,92% of the test cases passed. Also, 15,38% did not pass or had no results and 7,7% passed, but had a warning.

In TC-21, for example, using a larger number of rated products, interests and a higher p_l, its possible to see that the larger the p_l is, the longer it takes to run the algorithm overall. Another factor that could be observed was that the results were almost the same even with some different p_l values.

Besides that, some test cases with different product categories still unsuccessful, but its a smaller number of test cases than the cycle one due to a change to the algorithm that was made after the first cycle. Also, some general bugs from the App were found and fixed during this phase.

The Table 5 shows the results from the third cycle of tests. The third cycle contains tests of gray type. The objective of the third cycle was to test a higher quantity of "n" candidates products listed (w). The values of "n" were 50, 100, 200 or 300. And the quantities of interests and rated products were 1 or 2.

Table 6: Third cycle of tests.

TC-1	TC-2	TC-3	TC-4	TC-5	TC-6	TC-7	TC-8

To make the execution of these tests possible, it was necessary to adapt the pagination of products results. However, those tests about the "n" candidates products listed (w) were still inconclusive, more tests are necessary in the future to define.

After these tests, it was possible to define a value of p_l equal to 10. In this cycle, it was possible to observe again that most of the test cases with different product categories had only one result. That is because the main keyword is from the rated product that has no corresponding category with the interests. And that is much more specific than the case that both have the same categories. That will be improved in the future.

Moreover, the recommendations quality is above the average. And the time of the algorithm execution was an acceptable period considering the standards time of execution in other applications, the hardware and the complexity of it.

4 CONCLUSIONS

This work allowed to present the results of the previous paper, (de Paula Pereira et al., 2017), presented the entire theoretical part of the algorithm implemented in this paper and research review, which investigated possible gift recommendation solutions that take into account the user profile and among the solutions found, the one that best matches the context of this work is the COREL framework.

The COREL framework was customized to the Gift application context, which required the recommendation algorithm to run locally in the device and recommend gifts based in the user preferences. The proposed algorithm was tested through three cycles of gray and black box to verify and validate if it was working as expected and define some constants. A large number of improvements were made during this process and the results presented pointed goals were accomplished, the algorithm presented a good recommendation gifts and process performance.

REFERENCES

Buscapé (2016). *http://www.buscape.com.br*. Accessed on: November 26th 2016.

Company, B. (2016). *http://developer.buscape.com.br/portal*. Accessed on: November 26th 2016.

da Costa, R. P., de Paula Pereira, C., and Canedo, E. D. (2017). *Algoritmo de recomendação de presentes em dispositivos móveis*. Universidade de Brasília - Faculdade Gama. https://fga.unb.br/articles/0001/9506/tcc-gifter-caique-2.pdf.

de Braslia, U. C. (2016). *http://www.bepiducb.com.br/*. Distrito Federal DF, Brasil. Brazilian Education Programm for iOS Development.

de Paula Pereira, C., da Costa, R. P., and Canedo, E. D. (2017). *Mobile Gift Recommendation Algorithm*. 12th Iberian Conference on Information Systems and Technologies (CISTI).

de Souza, A. E. R. (2013). *Um modelo para recomendação de cursos de especialização baseado no perfil profissional do candidato. Dissertação (Mestrado) - Universidade Presbiteriana Mackenzie - So Paulo*.

do Nascimento, A. R., da Ssilva, B. F., and dos Santos, G. G. (2009). *E-commerce: O Melhor Caminho no Mercado Atual*.

Gama, R., André, N., Pereira, C., Almeida, L., and Pinto, P. (2011). *Algoritmo de Recomendação Baseado em Passeios Aleatórios num Grafo Bipartido*.

Lomadee (2016a). *http://developer.buscape.com.br/portal/lomadee/api-de-ofertas/introducao*. Accessed on: November 26th 2016.

Lomadee, O. A. (2016b). *http://developer.buscape.com.br/portal/lomadee/api-de-ofertas/recursos*. Accessed on: November 26th 2016.

Lomadee, P. (2016c). *http://developer.buscape.com.br/portal/lomadee*. Accessed on: November 26th 2016.

Mendes, R. (2016). Os números do mercado de e-commerce. Accessed on: November 26th 2016.

Qiu, J., Lin, Z., and Li, Y. (2015). *Predicting customer purchase behavior in the e-commerce context. Springer Science.*