ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

A Time Based Personalized Retail Recommender Strategy Using Data Mining

Prasanna Ravichandran[1], Shantanu Singh[1], Srinadh Dasaka[1], Vinaya B[1], A Rafega Beham[2]

B.E Student, Dept. of ISE, New Horizon College of Engineering, Bangalore, Karnataka, India[1]

Senior Assistant Professor, Dept. of ISE, New Horizon College of Engineering, Bangalore, Karnataka, India[2]

ABSTRACT: E-Commerce companies are growing fast, they concurred all the domains in the market. Recommendation algorithms are best known to be used by E-commerce companies that is the important reason for their success. Retails stores can't withstand in the competition, there is no efficient recommendation system used by the retail stores. E-Commerce platforms provide recommendations to the user based on their previous purchase. Considering time interval in recommending a product attracts the customer to buy more products. If it is implemented on retail stores it helps them to make more profit and sustain in the market.

KEYWORDS: E-Commerce, Recommendation, Time based, Retail stores, Filtering

I. INTRODUCTION

Recommender systems are software tools which uses different algorithms to suggest an item to the user. Such as what movie to watch and what to buy etc. So recommender systems helps a lot for the companies using it in many ways. The different recommending strategies can be categorized into content-based filtering recommendation, collaborative filtering based recommendation, association rule based recommendation and hybrid recommendation. Collaborative filtering predicts the user's favourite product by analysing the behaviour. It is one of the most promising algorithm used for recommender systems.

Since the retail stores are not using efficient recommendation systems they can't compete with the E-Commerce giants. Most of the people are not interested in going to the retail stores nowadays because of the attractiveness of the E-Commerce websites. The recommender algorithms used by the E-commerce uses the customer's previous purchase details and neighbourhood matchings to recommend a product to the customer, it is not efficient and the probability that the customer buys the recommended product is very less. So we have to consider a new way to predict the customer needs and recommend a product. It should add more accuracy to the previous algorithms.

In this paper we propose a time based recommendation technique which uses customer's purchase pattern in buying a product and recommends it on proper time. When we choose a proper attribute for predicting the recommendation time then it is a success.

II. RELATED WORK

Recommender systems have been applied to various domains such as book, music, groceries and movie recommendations. With the increasing amount of requirements, increasing amount of products are sold in E-Commerce websites and in retail stores. These recommender systems helps a lot in selling of products. Researchers have made a lot of progress in recommender systems Content based filtering and collaborative filtering methods are proposed by Lu et.al[3]. Collaborative filtering algorithm can be divided into two memory based and model based. It can be also divided into neighbour-based and model based. Although their names are different, the working and the content of classification algorithm are similar.

Antonio et.al suggested a prediction method of collaborative filtering recommendation for the ratings of user based on Bayesian probabilistic model. Fulizhang[1] proposed a time based recommendation, which uses collaborative filtering with time interval based recommendation techniques.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

The memory based collaborative filtering algorithm calculates the similarity between products and the user according to the existing datasets available and selects users or products that contain high similarity as neighbours of the target users. Calculating the rating of neighbours is used to predict the target user's preference for a product.
The model based collaborative filtering algorithm will make a model by learning the training dataset and uses the model to predict the unknown data. It also includes the clustering technique based collaborative filtering, probability-method based collaborative filtering and matrix decomposition based collaborative filtering etc.
The traditional collaborative filtering has more advantages that it can provide recommendation service for users under the circumstances of not considering the content of recommended items. The ratings of the related items are calculated usually in the process of recommendation, but the time sequence information is ignored by all. The time sequence information will improve the recommendation algorithm. It adds time sequence information into the existing recommendation model. It enables the model to learn the data changing overtime. As a result, the accuracy of the recommendation results would be improved. Fulizhang[1] proposed an improved recommendation system with time series adjustment. Which uses time interval in purchasing a product uses it for recommending the product next time.
In this paper we have used the algorithm defined by Fulizhang[1], which uses time sequence based information in the collaborative filtering. It increases the efficiency and correctness of the recommendation system. As it predicts the exact depletion time of a product and recommends it at proper time.

III. ALGORITHM DESIGN

In the fast moving world, it is hard for the customers to keep track of the quantity of the groceries remaining in their house. Sometimes they are not aware of the depletion of a product then they will be in critical situation. The shops can predict the depletion time of the product and recommends it to the customers. Customer with same number of family members can be grouped together as there are chances that they will use the products in same interval
Integrating time series information into the traditional product recommendation algorithm plays an important role in improving the accuracy of the product recommendation.

On the basis of time-sequence based collaborative filtering recommending algorithm combined with time characteristics of buying a product, this paper proposes a personalised time based retail recommender strategy based on time sequence inn buying a product. We adopt two characteristics factors inn our algorithm, the time sequence information in buying a product and product depletion time. Our algorithm takes these two factors a input and calculates the distance. Recommendation results are generated according to the distance values.
The motivation of our algorithm is as follows

A) Algorithm is easy to implement, it is simple to program. Traditionally used recommendation algorithm just suggests products based on customer's previous purchase. Sometimes customers do not like to see those. In this paper, two variables involved in this algorithm can be found in any time series based management system and the recommendation can be given without user participation, which is easier to be used by the store side.
B) The recommendation results conform to the customers buying pattern. Customers with the same number of family members are grouped together. So the time interval in buying a product will be same for those families. So it uses one families buying pattern to predict the recommendation time of another similar family.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

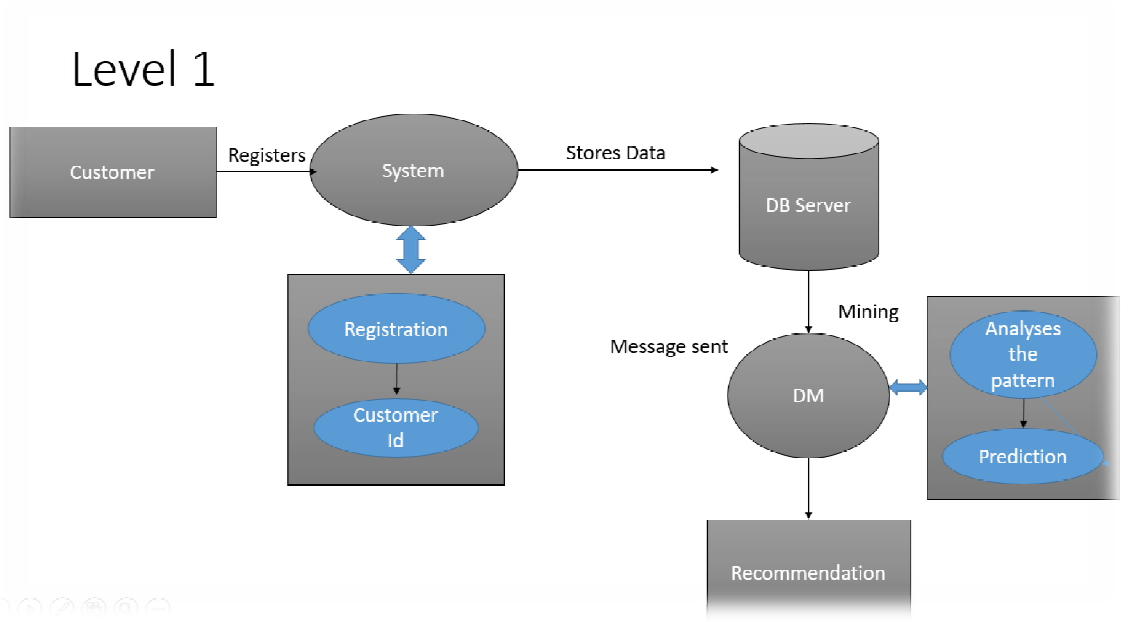
An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

Main contribution of our proposed product recommendation algorithm is that the recommendation results are based on the time pattern in buying a product. As it is important attribute in predicting the need of the customer.



a) Algorithm Description

When the customer buys a product from the retail stores the bill details including time of purchase and customer's family details are stored in the database. Our algorithm uses distance in time stamp of the time the product was bought by the user at different days. Two variables used in the calculation of the distance are the time sequence information in buying product and the depletion time of the product.

If the customers are grouped then it will be easy to find the time to recommend and what product to recommend.

The symbols used in the algorithm are

P_j – Product that customer buys from the store.

$L = (\{l_1\}, \{l_2\}....\{l_m\})$ is the set of users bought product P_j

m – Element number in set L

$l_i = \{(t_{i1},P_{i1}),(t_{i2},P_{i2}).....(t_{in},P_{in})\}$ & L is the customer record of buying product P_j .

t_{ij} – time the product P_j was bought

Our algorithm is described as follows: First,Algorithm looks up the records of readers who have bought product P_j and generate the set L. Then it calculates the time sequence information and the depletion time of the product P_j of different quantity.

The time sequence information is the difference between t_{i1} of P_j and t_{i2} of P_j and more. It calculates the time and day difference in buying the same product P_j for different days . It finds average time difference T_k of buying the product P_j.

$$T_j = \sum_{i=1}^{m} (t_{ij+1} - t_{ij})/N_j$$

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

$$\text{dis}(P_j,P_j) = \sqrt{Tj^2 + (N' - Nj)\text{^}2}$$

N_j – No.of times $P_j$appeared after the first time
N' – Max number of N_j

	Id	Product ID	Descriptic	Quantity	InvoiceDate	UnitPrice	CustomerID
2	536365	85123	Wheat	1	01-12-2010	2.55	17850
3	536365	71053	Basmati ri	1	01-12-2010	3.39	17850
4	536365	84406	Maida	1	01-12-2010	2.75	17850
5	536365	84029	Seasame (	1	01-12-2010	3.39	17850
6	536365	22749	Sunflowei	1	01-12-2010	3.39	17850
7	536365	22752	Urad dhal	1	01-12-2010	7.65	17850
8	536365	21730	Moong dh	1	01-12-2010	4.25	17850
9	536366	22222	Idly rice	1	01-12-2010	1.85	17850
10	536366	85123	Wheat	1	01-12-2010	1.85	17850
11	536367	71053	Basmati ri	2	01-12-2010	1.69	13047
49	536372	22749	Sunflowei	1	20-12-2010	1.85	17850
50	536372	22752	Urad dhal	1	20-12-2010	1.85	17850
51	536373	21730	Moong dh	1	20-12-2010	2.55	17850
52	536373	22222	Idly rice	1	20-12-2010	3.39	17850
53	536373	85123	Wheat	1	20-12-2010	2.75	17850
54	536373	71053	Basmati ri	1	20-12-2010	4.95	17850
55	536373	84406	Maida	1	20-12-2010	1.06	17850

The table consists of customers bill details. It consists of product's name and Id , date of purchase and customer. The customer having Id 17850 has bought the same products after 19 days of interval. So this interval and product's quantity are used to predict the next recommendation date.
Consider the product Maida as P_1 , t_{ij} is the time P_1 was bought and t_{ij+1} is time when P_1 was bought second time.
T_1 = ((20-12-2010 10:51) – (01-12-2010 8:26)) /1= 19 days
So the product maida is recommended to customer again after 18 days. Provided that the customer didn't come back again and purchase maida. This cycle continues, algorithm keeps on calculating the next recommendation date whenever the customer buys that product. If the customer changes the quantity of the product every time then an approximate date is chosen and recommendation is done.
The dis(p_j,p_j) is used to find the relation between the same product bought at different time. It is very useful if the customer buys the product in different quantity every time.

b) Processing Method of different quantity of product

The reason for inaccurate and inconsistent with the product information and solutions are summarized.

i) Customer may buy different quantity every time, it is difficult to analyse and predict the depletion time in this situation. We need to introduce statistical algorithm to solve this problem by finding an approximate date for sending recommendation.

ii) Product information input errors arise from the carelessness of cataloguing staff. So it is difficult to merge new products in the database. This kind of errors often appears in the starting stage of manual cataloguing and can only be dealt with manually.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

C) Cold start problem

It means the recommender system not able recommend product. When the customer goes to the shop for first time, the system doesn't have any of his purchasing pattern or database. So they can't precisely recommend a product to that customer. A new product in the store also can't be recommended as it's not sure that everyone will like that. This problem can be rectified by using existing user's data. The new customer's family details and purchase details are used to compare with others, if similar is found then that data is used for the recommendation purpose.

IV. EXPERIMENT AND ANALYSIS

Experiment datasets are all grocery store records from 2010 -2011

Experiment 1:Calculate the date of recommendation

We selected one customer and analysed his purchase pattern in buying some products. We found an interesting pattern in view of recommendation results. The customer had purchased similar products in same interval. So the recommender system analyses the date and quantity of the product bought each time and gave us one output date. The date it recommended was accurate, it correctly predicted the depletion time of the product in customer's house.

We selected a customer with customer Id 17850, our system analysed his purchasing pattern and gave us his previous bill details and profile details. Then his product details were analysed and one particular product is chosen sunflower oil, Id 22749.

So now the customer's last two purchasing dates are analysed and the difference is found, it will be in numeric type. This difference is then used to find an approximate date to recommend then product.

Customer 17850 had come to the store 4 times in 2 months, he bought sunflower oil in all 3 times and it's of same quantity.The difference in the dates are analysed, it is 19 days approx., so again the sunflower oil is recommended after 19 days of his last visit.

```
For the customer 17850 and  22749  product


@relation 'data4-weka.filters.unsupervised.instance.SubsetByExpression-EATT7 = 17850-weka.filters.unsupervised.instance.SubsetByExpression-EATT2 = 22749'

@attribute Id numeric
@attribute 'Product ID' numeric
@attribute Description {Wheat,'Basmati rice',Maida,'Seasame oil','Sunflower oil','Urad dhal','Moong dhal','Idly rice',rice,sugar,salt}
@attribute Quantity numeric
@attribute InvoiceDate {01-12-2010,08-12-2010,18-12-2010,20-12-2010,23-12-2010,24-12-2010,25-12-2010,26-12-2010,27-12-2010,01-01-2011,02-01-2011,03-01-2011
@attribute UnitPrice numeric
@attribute CustomerID numeric

@data
536365,22749,'Sunflower oil',1,01-12-2010,3.39,17850
536372,22749,'Sunflower oil',1,20-12-2010,1.85,17850
536396,22749,'Sunflower oil',1,11-01-2011,1.06,17850
536409,22749,'Sunflower oil',1,30-01-2011,4.25,17850
01-12-2010
Days: 19
20-12-2010
Days: 22
11-01-2011
Days: 19
Possible interval for recommendation is :  19
```

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.5, Special Issue 5, June 2017

8th One Day National Conference on Innovation and Research in Information Technology (IRIT- 2017)

Organized by

Departments of ISE, CSE & MCA, New Horizon College of Engineering, Bengaluru, Karnataka 560103, India

V. CONCLUSION

This paper has proposed an efficient method of product recommendation using time sequence based information. This increases the accuracy of the algorithm, a product is recommended at proper time, and there is a high probability that it will be accurate. The dataset used also has values of different customers which is useful for the experimentation and for training the software. This time based recommendation can be used in many places like in vehicle service stations, Petrol stations etc. None of the traditional recommender system uses the time sequence information, it is critical but important attribute to concluded while predicting.

REFERENCES

[1]FULI ZHANG "A Personalized Time-Sequence-Based Book Recommendation Algorithm for Digital Libraries", IEEE, date of publication May 9, 2016

[2] Achin Jain, Vanita Jain and Nidhi Kapoor "A LITERATURE SURVEY ON RECOMMENDATION SYSTEM BASED ON SENTIMENTAL ANALYSIS", An January 2016, BharatiVidyapeeth College of Engineering, New Delhi

[3] Z. Lu, Z. Dou, J. Lian, X. Xie, and Q. Yang, ``Content-based collaborative filtering for news topic recommendation," in Proc. AAAI, 2015,pp. 217_223