Iraqi Journal of Science

ISSN: 0067-2904

Design recommendation system in e-commerce site

Ayad R. Abbas, Shaymaa Ashor*

Department of Computer Science , University of Technology , Baghdad , Iraq.

Abstract

In recent years it has spread the used of e-commerce sites quite dramatically. Thus, these sites have become display huge number of diverse products. It became difficulty for the customer to choose what he/she wants from this product. The recommender systems are used to help customers to finding the desired product of their interests and proved to be an important solution to information overload problem.

This paper, designed a recommendation system based on content, which is usually textual description. Furthermore, the proposed system uses cosine similarity function to find the similarities among the characteristics of various products, and nominate a suitable product closer to customer satisfaction. The experimental result shows that the proposed system can provide suitable product with accuracy up to 95%.

Keywords: E-commerce, Personalized recommendation system, Content-based, cosine similarity.

تصميم نظام توصية في موقع تجارة الكتروني

اياد عباس روضان ، شيماء عاشور حمداوي*

قسم علوم الحاسبات ، الجامعة التكنولوجية ، بغداد ، العراق.

الخلاصة

في السنوات الأخيرة انتشر استخدام مواقع التجارة الالكترونية بشكل كبير جدا وأصبحت هذه المواقع تعرض عدد كبير جدا من المنتجات المتنوعة. وأصبح من الصعب على الزبائن اختيار حاجاتهم من المنتجات أنظمة التوصية تستخدم لمساعدة الزبائن على إيجاد المنتجات المناسبة لاحتياجاتهم وتقدم حل مهم لمشكلة الزيادة في المعلومات.

في هذه البحث تم تصميم نظام توصية بالاعتماد على محتوى المنتجات، والذي عادة يستخدم الوصف النصي. علاوة على ذلك ، أن النظام المقترح يستخدم دالة الجيب تمام لإيجاد التشابه بين صفات المنتجات المختلفة ، ثم يتم ترشيح المنتجات المناسبة الأقرب لرضا الزبون. أظهرت التجارب أن النظام المقترح ممكن أن يزود الزبون بمنتجات مناسبة مع دقة تصل إلى 95.%

1. Introduction:

With the widespread use of the internet all over the world, spread using e commerce sites for the sale products to customers and become companies compete to provide goods which meet the desired customers. Personalized recommender systems is used to help customers to finding the product of their interests and proved to be an important solution to information overload problem and increase

* Email: shaymaashor1111@gmail.com

sealing of companies at the same time by increasing the loyalty of customers to those companies by making the procurement process is completed with the fewest number of steps and complexity [1-2].
There are six types of systems Recommendation:

1. **Content-based**: this type is used to suggestion products similar to one impressed customer in the past, usually used textual description or structural of product to find suggestion products [3-4]. The advantage of this type that it does not need the opinions of the rest of the customers in order to take a decision filtration products as well as Capable of handling the new product problem that has not been rated by any customer and it can deal with the problem of cold start, the disadvantage of this type Nominate only products similar to that chosen by the customer during the search [1-2,4-6].
2. **Collaborative filtering(CF)**: this type used to suggestion products automatically to the active costumer by measured rating similarity with other costumers or product ,which can be divided into two approaches :user-based CF and item-based CF approach. The advantage of this type is the ability to diversify by offering products discovered fortuitously, but suffers from the problem of Scalability, spared data and Cold-start[1-2,6].
3. **Demographic:** information like age, Gender, Employee and language belonging to a particular customer is used in this type to suggestion personalized products. this type is not needed extra knowledge or history of ratings , so demographic recommendation systems suffers from the problem of limited accuracy, to improve the accuracy More information such as textual reviews is required of the recommendation[1,3,8].
4. **Knowledge-based:** This kind based on explicit knowledge about how certain product features meet costumer's requirements and preferences. This system used similarity function to measure how much the recommendation match costumer's requirements.[1,7]
5. **Community-based:** Depending on the views of the customer's friends the products are recommended in this type [1].
6. **Hybrid recommender systems:** this type combining two or more of the previous types and it is capable of solved problem of the previous types, but it suffers from the cost and complexity [1-2].

There are many applications that use different types of these systems like Amazon.com, Netflix, Movelins and facbook ect. In E-commerce sites, in particular, information provided to the customer before and during the procurement process greatly affect to the decision procure of the costumer, So the process of gathering information about the customer and the product and then display the results to the customer must be efficient and convenient for the customer with the lowest number of steps.

Most recommendation systems need to rate the customer to particular product in order to recommended other products, rating done in two ways explicitly or implicitly. In explicitly feedback rating is directly done by the customer, for example, in Amazon.com the numerical ratings such as the 1-5 stars provided in the book recommender. While implicitly feedback rating is used when explicitly feedback did not fond It depends on browsing history or click to rating the products, this lead to eliminates the cost of providing the customer rated [1,9].

The proposed system used content based recommendation system and used implicitly feedback rating to collecting information to suggestion products similarly to that product selected by the customer in the past with high relevant Without asking the customer registration or any personal information.

This paper organized as fellows, section 2 consist of related works**,** section 3 consist of details of the proposed system, section 4 consist of conclusion

2. Related works

Many studies have been submitted to address the problems of recommendation systems in several respects such as:

1. Item-to-Item Collaborative Filtering is used to suggest products in real time similar to other product purchased or rating by a particular customer by using cosine similarity function. This technique is applied in Amazon Customers who bought [10].
2. In a digital library to help library users in finding most relevant research papers over a large volume of research papers, recommender systems is proposed based on adopted content-based

filtering technique. The presented algorithm employs both TF-IDF weighing and cosine similarity measure to suggestion recommendation systems based the users' query to the intended users[11].

3. For sparse data dynamic personalized recommendation algorithm proposed in this paper. one prediction with more rating data is utilized during each feature by connecting more neighboring ratings in customer and product profiles, to illustrate the preference information based on TSA technique a set of dynamic features are designed, and finally adaptively weighting the features in different phases of interest using information to made recommendation list[12].
4. **[13]** The authors worked on recommendation systems for music data by using genetic algorithm and the content based filtering technique. The system tries to nominate items satisfy customer based on the analysis of the data used by that customer.
5. The system design to recommended product based on customer's views. Which is called expert, expert is a person who bought a product previously and has an expert about the quality of this product .The view of this person as possible to help the customer to choose the suitable product [14].

3. Proposed System

The proposed system used content-based recommendation system because it has the ability to dynamically deal with the products that described by texts without need to rating the product by customer ,and it can deal with the problem of cold start ,new user ,new product and scalability problem.

3.1. Products description:

The dataset are utilized as a part of this paper consists of 320 office products [15]. These product which extracted from the site have been arranged in a table inside a private database are decrypting by a set of attributes. These attributes are:

- ID of product.
- Title of product.
- Type of product.
- Price.
- Image: image name in the data base.
- Brand: Brand name for the product.
- Categories: It contains a text description of the product.

3.2.The processing steps:

The processing steps to recommend a list of product for active user are:

Step 1: Costumer Select particular product form the list of product as shown in Figure-1

.

Figure 1

Step 2: Extraction of all products that match this product by type and displayed to the customer as shown in Figure- 2

Figure 2

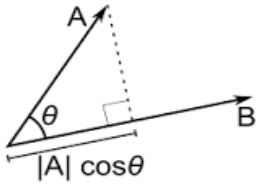
Step3: When the costumer select one of these products, they system display the details of this product, at the same time nominate products the customer might wish to purchase as shown in Figure-3.

Figure 3

Step 4: used ID of this product to extract the categories, and then convert it to the array of token(x).
Step 5: For each type of products extracted the categories and convert it to the array of token(y_i).
i= 1…n, n number of product type.
Step 6: integrates array(x) with array (y_i) in unique array(k) without repeating the type of words.
Step 7: generate vector (A) by searching for each value of the array(x) within the array(k) if they value do not exist put one in vector(A) else put zero in vector (A).
Step 8: repeat step7 for each type of product to generate vector (B_i).
Step 9: using cosine similarity to measure the similarity between vector (A) and vector(B_i).If the angle between vectors close to zero as a result of the similarity is equal to one. This indicates that the amount of similarity is very high and If the angle between the vectors close to 90 as a result of the similarity is equal to zero. This indicates that the amount of similarity is very low, As it is shown the following example:
If the active user select the product that type Appointment Books & Planners,
The system used id to extract the category.
Category for select product is " Office Products, Office & School Supplies, Calendars, Planners & Personal Organizers, Appointment Books & Planners "
Which Convert to array of token (x)= " Office ,Products, & ,School, Supplies, Calendars, Planners , Personal, Organizers, Appointment, Books ". and then extract
Other product Category = "Office Products, Office & School Supplies, Education & Crafts, Arts & Crafts Supplies, Classroom Decorations"
convert to array of token (y_i)= " Office ,Products, &, School, Supplies, Education , Crafts, Arts , Supplies, Classroom ,Decorations "
Each array represents the vector as follows:
Vector A= x (1,1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0)

Vector B= y (1,1,1,1,1,1,0,0,0,0,0,0,1,1,1,1,1)
Then applied they function of similarity as shown in equation 1.

$$sim(\vec{A},\vec{B}) = \cos(\emptyset) = \frac{\vec{A}\cdot\vec{B}}{\|\vec{A}\|*\|\vec{B}\|} \qquad \text{……..(1).}$$



The following table shows the experimental results of application the similarity equation between the product of type (Appointment Books & Planners) and the other type of products.

Table 1- Show the results of using the cosine similarity function

Product name	**Similarity value**
Appointment Books & Planners	**1**
Classroom Decorations	0.52
Class Records & Lesson Books	0.58
Paper	0.65
Wide-format Paper	0.57
Copy & Multipurpose Paper	0.57
Inkjet Printer Paper	0.23
Photo Paper	0.54
Dry Erase Boards	0.48
Impact & Dot Matrix Printer Ribbons	0.26
Pins & Tacks	0.52
Rubber Bands	0.52
Correction Tape	0.54
Erasable Markers	0.52
Mechanical Pencil Leads	0.5
Wooden Colored Pencils	0.52
Ballpoint Pens	0.52
Compasses	0.52
Protractors	0.45
Printers	0.43
Conference Room Tables	0.38
Calculators Basic	0.54
Laser Printers	0.43
Lateral File Cabinets	0.32
Mobile File Cabinets	0.34
Vertical File Cabinets	0.34
Racks & Displays	0.27
Shelving & Storage	0.27
Storage Cabinets	0.27
Lecterns & Podiums	0.28
Mail Carts	0.30
Desk Chairs	0.30
Chairs & Sofas,Stools	0.30
Chairs & Sofas,Task Chairs	0.30
Computer Workstations	0.38
Office Desks	0.40
Bulletin Boards	0.51
Inkjet Printer Ink	0.30
Pencil Sharpeners	0.48
Mechanical Pencils	0.48
Desktop Staplers	0.48
Liquid White Glues	0.48
Store Sign Holders	0.52

Step 10: If the result of similarity >= 0.48 and =< 0.95, the product add to the recommendation list.

Evaluation metrics:

Recall, precision, and accuracy functions used to evaluate the efficiency of proposed system [1,16] as shown in Table- 2 **.**

Table- 2

	The system can recommend these product	The system cannot recommend these products
predictable products	True positive TP	False negative FN
Not a predictable products	False positive FP	True negative TN

RECALL Is the percentage of related records that have been recovered to the entirety number of related records in the database.

Recall = $\frac{TP}{TP+FN}$

PRECISION Is the percentage of related records that have been recovered to the entirety number of unrelated and related records that have been recovered.

PRECISION = $\frac{TP}{TP+FP}$

Accuracy Is the percentage of products that must be predictable by the system to entirety number of all products recommended

ACCURACY = $\frac{TP+TN}{TP+TN+FP+FN}$

Results are shown in the following Table- 3

Table 3- the results of recall, Precision and Accuracy

Product name	Recall	Precision	Accuracy
Appointment Books & Planners	**1**	**0.8**	**0.9**
Classroom Decorations	**1**	**1**	**1**
Class Records & Lesson Books	**1**	**0.8**	**0.9**
Paper	**0.6**	**1**	**0.76**
Wide-format Paper	**0.83**	**0.83**	**0.95**
Copy & Multipurpose Paper	**0.8**	**1**	**0.95**
Inkjet Printer Paper	**0.9**	**1**	**0.97**
Photo Paper	**0.83**	**1**	**0.97**
Dry Erase Boards	**1**	**1**	**1**
Impact & Dot Matrix Printer Ribbons	**1**	**0.7**	**0.95**
Pins & Tacks	**1**	**1**	**1**
Rubber Bands	**1**	**1**	**1**
Correction Tape	**1**	**1**	**1**
Erasable Markers	**1**	**1**	**1**
Mechanical Pencil Leads	**1**	**0.9**	**0.95**
Wooden Colored Pencils	**1**	**1**	**1**
Ballpoint Pens	**1**	**1**	**1**
Compasses	**1**	**1**	**1**
Protractors	**1**	**1**	**1**
Printers	**0.43**	**1**	**0.8**
Conference Room Tables	**1**	**0.9**	**0.95**
Calculators Basic	**0.92**	**1**	**0.95**
Laser Printers	**0.7**	**1**	**0.92**
Lateral File Cabinets	**1**	**1**	**1**
Mobile File Cabinets	**1**	**1**	**1**
Vertical File Cabinets	**1**	**1**	**1**
Racks & Displays	**1**	**1**	**1**
Shelving & Storage	**1**	**0.92**	**0.97**

Storage Cabinets	**1**	**0.92**	**0.97**
Lecterns & Podiums	**1**	**1**	**1**
Mail Carts	**0.86**	**1**	**0.95**
Desk Chairs	**0.86**	**1**	**0.95**
Chairs & Sofas,Stools	**0.86**	**1**	**0.95**
Chairs & Sofas,Task Chairs	**0.86**	**1**	**0.95**
Computer Workstations	**0.86**	**1**	**0.95**
Office Desks	**0.72**	**1**	**0.88**
Bulletin Boards	**1**	**1**	**1**
Inkjet Printer Ink	**1**	**0.85**	**0.97**
Pencil Sharpeners	**1**	**0.85**	**0.92**
Mechanical Pencils	**1**	**0.85**	**0.92**
Desktop Staplers	**0.95**	**1**	**0.97**
Liquid White Glues	**1**	**0.7**	**0.85**
Store Sign Holders	**1**	**1**	**1**
Average	**0.93**	**0.95**	**0.95**

Conclusion:

In this proposed system used content-based recommendation system because it has the ability to dynamically deal with the products that described by texts without need to rating the product by customer, and it can deal with the problem of cold start, new user, new product and scalability problem. This system is characterized by high efficiency and accuracy in recommended products as shown results that appeared.

References:

1. Ricci.F, Rokach.L and Shapira.B. **2011**. *Introduction to Recommender Systems Handbook*. Springer, pages 1-200.
2. Kumar.P.N, and Reddy.V. **2014**. A Survey on Recommender Systems (RSS) and Its Applications. *International Journal of Innovative Research in Computer and Communication Engineering(IJIRCCE)*, Vol. 2, Issue 8, August.
3. Kazienko.P and Kolodziejski.P. **2006**. Personalized Integration of Recommendation Methods for E-commerce. *International Journal of Computer Science & Applications.* Vol. 3 Issue 3, pp12-26.
4. Hong.W, Li.L. and Li.T. **2012**. Product recommendation with temporal dynamics. *Expert Systems with Applications* 39,12398-12406.
5. Chaudhar.M and Patil.S. **2015**. A Survey for Personalized Item based Recommendation System. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS).* Volume 4, Issue 1, January-February.
6. Gong.S. **2010.** A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item Clustering. *Journal of software*, Vol. 5, No. 7, JULY.
7. Jun.H, Fan.Y, Jun.T and Shuo.C. **2013.** A comparative study of the recommendation Algorithm Applicator in the network of educational resources Platform**.** International Conference on Education Technology and Information System (ICETIS).
8. Wang.Y, Chan.S.C and Ngai.G. **2012.** Applicability of Demographic Recommender System to Tourist Attractions: A Case Study on Trip Advisor. IEEE/WIC/ACM International Conferences on (Volume:3)
9. Zhang.B, Guan.Y , Sun.H , Liu.Q and Kong.J. **2010.** Survey of user behaviors as implicit feedback .International Conference on Computer, Mechatronics, Control and Electronic Engineering (Volume:6).
10. Linden.G, Smith.B, and York.J. **2003.**Amazon.com Recommendations Item-to-Item Collaborative Filtering. *Published by the IEEE Computer Society*.
11. Philip.S, Shola P.B and John.A.O. **2014.** Application of Content-Based Approach in Research Paper Recommendation System for a Digital Library. *(IJACSA) International Journal of Advanced Computer Science and Applications*, Vol. 5, No. 10.

12. Prasanth.B and Latha.R. **2015.** Dynamic Personalized Recommendation Algorithm on Sparse Data. *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE).* Vol. 3, Issue 2, February.
13. Sultana.T, Sharma.N, Aher.S and Kinikar.N.P. **2015.** Recommendation System for Music File. *(IJCSIT) International Journal of Computer Science and Information Technologies.* Vol. 6 (1).
14. Vaishnavi.S and Karthik.S. **2013.** Design of Recommender System Based On Customer Reviews. *IJRET: International Journal of Research in Engineering and Technology.* Volume: 02 Issue: 11 | Nov.
15. (http://jmcauley.ucsd.edu data/amazon).
16. Lopes.P and Roy.B. **2015.** Dynamic Recommendation system using wab usage mining for E-commerce users. *ELSEVIER*, Procedia Computer Science 45. 60-69.