Middle-East Journal of Scientific Research 13 (6): 823-828, 2013
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2013.13.6.2497

Recommender System Based on Association of Complementary and Similarity in Electronic Market

Majid Khalaji and S. Javad Mirabedini

Department of Engineering Software, Arak Azad University of Iran

Abstract: Due to developments of Information Technology, most of companies and E-shops are looking for selling their products by the Web. These companies increasingly try to sellproducts and promote their selling strategies by personalization. In this paper, we try to design a Recommender System using association of complementary and similarity among goods and commodities and offer the best goods based on personal needs and interests. We will use of Ontology that can calculate the degree of complementary, the set of complementary products and the similarity, then offer to users. In this paper, we identify two algorithm, CSPAPT and CSPOPT. They have been offer better results in comparison with the algorithm of rules; also they don't have cool start and scalable problems in Recommender Systems.

Key words: Recommender Systems · Collaborative Filtering · Ontology · Semantic Complementary · Semantic Similarity · Association Rule

INTRODUCTION

Today, the world wide web is a great place of digital documents considering the developments of information technology.With increasing information references, numbers of choices have been become more and more to find required goods.Also, it is difficult to maintain this information and this problem is related to the developments of information technology and is known as the expensive overload. Personalization is one of the solutions.Personalize systems benefit from a Recommender System and try to settle dynamic pages based on personal interests. Recommender Systems are important research references. The first related paper was presented in 1990 and was about collaborative filtering [1-2]. In recent decades, researches have been done that led to the developments of new methods for Recommender Systems. There are examples of applications such as products of Amazon site and Video sites.Recommender Systems have a significant effect on the success of websites as well as on the uncontrolled growth of electronic market and web-based systems. In information technology, the success of a website depends on customers and visitors attractions and definition of users` needs are necessary for improving applications of a website. Recommender Systems try to predict interests and needs of users using of data collections and offer a list of user needs. These Recommender Systems have many problems and led to some problems in great websites about offering goods and commodities to users. The relationships of goods have important function in designing of Recommender Systems. The most important relationships can be named complementary association and similarity.

In the other hand, ontology is used to determine the relationships among contents. Considering the structure of ontology, we can use it for determining the relationship of complementary association, similarity and other associations. Many problems of algorithms can be solved by the association between ontology and commodity.

The purpose of this paper is to design a Recommender System that can offer personal needs and interests. The offers are not the only response to the user needs, but also they pay so attention to user interests and tastes. The designed systems should have a good speed and accuracy that led to customers` satisfaction. Also, the other purpose of this paper is to solve some problems, such as cold start and scalable. These problems are in many Recommender Systems.

Corresponding Author: Majid Khalaji, Department of Engineering Software, Arak Azad University of Iran.

The most important achievements of this papercan be outlined as follows:

- Implementation and evaluation of the Recommender Systems using ontology and similarity association among goods.
- Offer goods to users using Recommender Systems in websites and sell through E-shops and usersatisfaction.
- Presentation a new model using data structures to enhance the offers of personalize systems and implementation and evaluation of offering method in this model.

In section 2, we identify Recommender Systems; In section 3, we explain complementary procedure and define subsections and the features of complementary and the way of calculation. In section 4, we design Recommender System based on complementary goods. In section 5, we evaluate the designed Recommender System and finally in section 6, we get a conclusion.

Recommender System: Web producers raise personalize web for better communication between customers and producers. Although web personalizationis not essential part of business, It has many applications in market. The purpose of personalization system is considering to user needs in the web. After personalization, Recommender Systems was presented according to user needs and tastes. Recommender Systems are divided into 3 categories based on the structure [3]:

- Rule – based systems: In rule – based systems, decisions were given based on the rules extracting automatically or manually through user information. The purpose of this system is discovering elements that effect on user preferences in choosing a product. ISCREEN is one of the rule – based systems. It`s used for filtering text messages by manually produced rules of users.
- Content – based filtering: This algorithm presents the offers based on the items that a user has sold before. Using the methods like Machine Learning and buying history identify to be attractive and not to be attractive of goods for customers. [4-5]. In these systems, items were offered based on content associations and user interests.
- Collaborative Filtering (CF): These systems offer the products based on behavior similarities and user application patterns. They perform statistical analysis by using data mining in databases, monitoring user behaviors, rating goods and buying history.
- Collaborative filtering is classified in two categories based on users, items and how it is planned [6-7].
- In the first classification, they find a user system that would be similar with the target system, then offer the goods to target users that is more attractive to users. In the second classification, the system finds the items that would be similar with the higher cost items and then offer to user and recognize the similarities of items based on given rates instead of finding the neighbor target users.

Recommender Systems have been designed based on concept associations.They pay too attention to user interests. But there is another considered factor that is user needs. For this reason, The similarity cannot recognize user needs. Therefore, another association is required compleentary association.

Complementary has many applications in economic science [8]. The complementary association of products means that the products costumed altogether like hamburgers and hamburger buns. In many application programs of Electronic Market, these kinds of associations are more beneficial than other associations. For example, offering a similar machine or bicycle to someone who buys a Mercedes Benz is not too beneficial.But offering a CD – player or an alarm system is more beneficial. After investigating these associations, Recommender Systems are designed using suitable associations correctly. In a designed system, firstly it`s offered complementary products to users, then for increasing accuracy of a system and satisfaction of user should be added by user taste about suggestions. It`s required to similarity. At last, Recommender Systems is designed for offering complementary products based on user tastes. Both complementary and similarity are used in a system that is going to be designed.

Complementary can be calculated based on is-a association of products and helping to a new property called need. In fact, we use ontology.

The purpose of this paper is to extract user needs through product associations and relationships. Product similarity is beneficial for determining user interests, but it cannot determine user needs. Therefore, to a relationship is required that is called Product Complementary in economy.

Economically, product complementary is a need product for using better the main product.For example, product A is complement of the product B. If it`s required to A for using better product B. The examples of product complement are hamburgers and hamburger buns.

In this paper, semantic procedure and semantic complementary are used. In this procedure, products are modeled by OWL language and extracted product complementary through ontology, then Recommender System is planned based on product complementary. Similarity illustrates the similarity between two products considering the associations of products. But complementary is another association that is different from similarity completely. For example, suppose the association between car and gasoline. In other word, suppose the relationship between car and bicycle.Car and bicycle have more similarities, but gasoline is more related to car (gasoline is complementary of car). It means that using of a car requires gasoline. [9-10].

Complementary Procedure

Product Complementary Property: Complementary is not bilateral relations.It means if A is complement B, we cannot say that B is complement A. For example, we need a video – player for using of a video - game; but we don`t need a video - game for using of a video – player. So, we can get the conclusion that the video – game is complementary the video – player.

Complementary has a transitive relation, it means if A is complementary B and B is complementary C, so A is complementary C.

Determining Complementary Degree: First, products should be divided based on standard UNSPSC. The standard indicates Is-a property in all items. In this step, the classes and the associations are determined. Every class indicates that it comes from ontology tags based on OWL language. A part of catalogue is related to fuel service that`s planned based on standard UNSPSC. See Figure1.

Classes have a property called need. It has an important function in calculating complementary algorithm degree. This property would be OWL in ontology`s language using property definitions. For example, a user needs a paper and an eraser for a pencil usage, so the need property includes an eraser and a paper. Com(A,B) illustrates complementary degree A and B, that is between 0 and 1. The closer to 1, the more need to B.

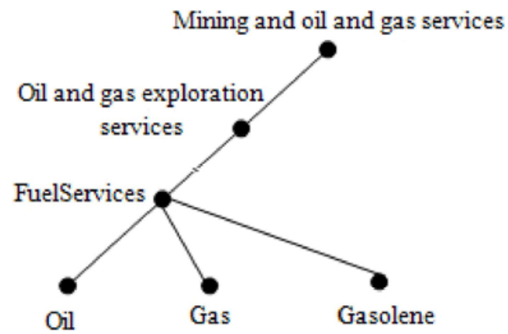

Fig. 1: Fuel Service Catalog

For determining Com(A,B), we require product associations and need properties. If B is one of the needs of A, so Com(A,B) is 1; if not we use coefficient k and the algorithm with B. This algorithm continues with the high class of B. If high class B is a member of the need product, so B is complementary A using k. If high class B is not a member of need A and high class B is not root, this procedure repeats with higher class B. The more distance between B and need A in catalogue, the less amount of Com(A,B) and if B and higher classes are not members of need A. So, the amount of Com(A,B) is equals to 0. The formula 1 indicates that:

$$com(A,B)=\begin{cases}1/\mathrm{k}\times \mathrm{com(A,\ B.H)} & \mathrm{B}\notin \mathrm{A.N}\\ 1 & \mathrm{B}\in \mathrm{A.N}\\ 0 & \mathrm{B.H}=\Phi\end{cases}$$

In the above formula, H is high class product. B.H is high class B. N is need property, that A.N indicates need A. k is variable coefficient that shows the increasing speed rate of complementary and equals 2 [11].

Example 1: Complementary Degree of Hamburger and Hamburger Buns.

First, draw a classification figure, then determine the products that are the amount of need hamburger and hamburger buns.

Burger.N = {Burger buns, Vegetable, Sause}
Com(Burger, Burger buns) = 1 Burgerbuns ϵ Burger.N

To Design a Recommender System: Recommender Systems are divided into 3 groups based on content filtering, correlative filtering and combined filtering. The designed system of this paper is not offered for increasing the algorithm accuracy like previous similar products but it is offered for combined complementary and similarity of user files. Also, it is used to concept method to determine complementary products for solving some limitations and problems such as first user, cold start and scalable. The whole purpose of this paper is to design a Recommender System for solving many problems of systems with the suitable accuracy and speed.

In these systems, users are divided into 2 classes. The first class is the users who have purchased the products before and the second class is the users who didn`t ever purchased the products from the market.

The system offers two classes of complementary and similarity products. To users First, while a user enters to a shop, the system offers some suitable products based on the user buying background. But the second group enters to the shop, the system cannot offer any product; Because, There is not any background. However along with a shopping, the system determines user need properties and offers suitable products.For this reason, complementary products are divided into two categories:
CSPAPT: Complementary and Similarity Product After Purchased Time
CSPOPT: Complementary and Similarity Product Over Purchased Time

CSPAPT Algorithm: When a user enters to a shop for the next time, the system offers the products based on CSPAPT algorithm. This algorithm checks previous purchases in the exact time that a user enters. Then it offers complementary and similarity products to previous products and it clearly determines colors, brands and sizes of the products. For instance, it offers a complement eraser to a user the buying a pencil. Considering the user who buys the pencil with identified color, size and brands, the system offers an eraser complement and similar with the pencil properties.

In this algorithm, the need products (complementary products) are determined then it offers the products which have the most similarity with the user interests and tastes. For example, the favorite color of a customer is blue and usually he/she buys office supplies with the brand X, the middle size and the quality Y. Therefore, the system offers to the customer who buys a pencil, a blue, brand X, quality Y and middle size eraser or it offers an eraser that most of customers buy it.

CSPOPT Algorithm: When a user is buying a product, CSPOPT algorithm offers the complementary and similarity products. When a user enters to a shop and

level4

id	name	is-a
1-1-2-1	motars	1-1-2
1-1-3-1	cement	1-1-3
1-1-3-2	lime	1-1-3
1-2-2-1	Sandlime brick	1-2-2
1-2-3-1	concrete tiles or flagstones	1-2-3
1-2-3-2	head stones	1-2-3
1-2-3-3	stone tiles or flagstones	1-2-3
1-2-3-4	ceramic tiles or flagstones	1-2-3
1-3-1-1	spout	1-3-1
1-3-1-2	sanitary ware	1-3-1
1-3-1-3	sanitary china	1-3-1
1-3-1-4	sink	1-3-1
1-4-1-1	asphalt	1-4-1
1-4-1-2	guny	1-4-1
3-1-1-1	gas burner	3-1-1

Fig. 2: The shop`s products classification

chooses some products and adds them to a basket, the system checks the basket and offers complementary and similarity products. For example, when a user adds a hamburger to a basket, the system checks the basket and offer the complementary hamburger (hamburger buns). If the hamburger is the best quality, the system offers the best hamburger buns.

The Recommender System works as a salesperson and tries to introduce the complementary and similarity products with purchased products to customers for more satisfaction and selling.

Evaluation: In this paper, a dataset of a building material company is used. The data has been collected from May 10, 2005 to August 01, 2006, based on invoices of the company that exists as a access database. The data consists of 2266 customers, 2581 products and 21662 invoices for simplifying to design and execute. the products are divided into 18 main classes, according to Figure 2 including, bricks, sand, cement, health products, Chinese products, Asphalt, sacks, gas, hood and so on.

It requires ontology for calculating complementary degree of both products. Product properties are in a table. The table is as a product`s ontology. The main column of this table is following as:

Id : Code of every product based on UNSPSC standard
Name : Name of every product
Is-a : High class of every product
Need : Need property of a product

In this issue, evaluation metric F1 is used. First, the purchased products are divided into two groups. The first group is Train Set and the other group is Hit Set. These sets are choose randomly. First, the given algorithm is executed on the train set and is called Top-N. Then, this products are compared with the test set and shared with both test and Top-N called Hit Set. Finally, an accurate percent of algorithm is determined after giving the test set, train set and hit set using evaluation metric.

$$F1 = 2* \frac{Size\ of\ Hit\ Set}{Size\ of\ Test\ Set + Size\ of\ Top-N\ Set}$$

We consider the mean of given F1s from all users as an accurate algorithm.

The first step of the system evaluation is to calculate complementary degree of products. In the second step, test set and train set are determined. As this sequence, 80% products exist in train set and the rest 20% products exist in test set randomly. In the third step, Top-N set is produced. The algorithm is executed on the train set and N products are produced for offering to users, the N products are the highest complementary degree with train set. In the last step, F1 is calculated for all users after producing train set, test set and hit set and the mean of F1 is considered as the final F1. We have done the algorithm with different sets, test and train sets 10 times and the result shows the algorithms CSPOPT and CSPAPT.

For evaluating the algorithm CSPAPT, the invoices of products are divided into tran and test sets and complementary product sets are classified based on color,size and quality. Then, identified complementary N product and Top-N set are made. The amount of F1 are calculated by test, Top-N and Hit sets. The final results have been shown in figure 3. In this diagram, the amount of F1 varies between 0/41 to 0/44.

For evaluating algorithm CSPOPT, the purchased products are classified based on the purchased date. In these algorithms, the purchased products are divided into train and test sets during a day. A complementary product set is classified based on color, size, quality and F1 are calculated. The final amounts of F1 is shown in Figure 4 and varied between 0/48 and 0/55.

Association Rules are known as algorithms that discover and find the rules and the relations among a big set of data. Discovering and finding the relations, rules of records and data in big databases of companies can be affected on better and attract manager decisions in different field such as catalogue designs, more and

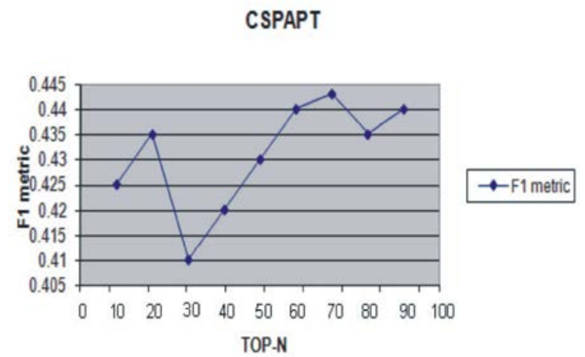

Fig. 3: Diagram F1 based on variable amount of N in algorithm CAPAPT

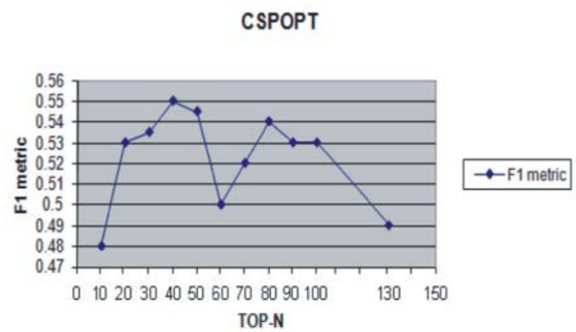

Fig. 4: Diagram of F1 based on the variable amounts of N in algorithm CSPOPT

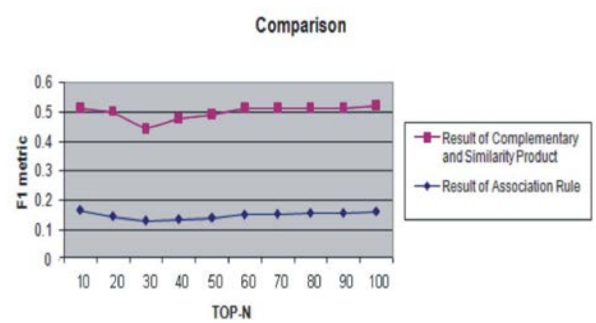

Fig. 5: Comparison F1 in association rules by CSPAPT

affective selling and various business and marketing policy. According to association rule algorithm, the Final amount of F1 is between 0/13 and 0/17.

Accuracy and result of CSPAPT and CSPOPT is more than association rules. Comparing with F1 through CSPAPT and association rules, it is clear that the results of association rules are poor. The F1 amount illustrates the amount of accuracy and system operations and executions. Also, it`s determined considering to algorithm comparisons. Recommender algorithms have better accuracy to association rules. The results is shown at Figure 5.

The other important criteria is the running time of algorithm. In the other word, the time is takento produce recommender products (Top – N sets). If an algorithm has a suitable and good accuracy and the time is too long, it leads to dissatisfactions. For this reason, the speed is very important point in E –shops.

CSPAPT and CSPOPT algorithms have more accuracy and speed to compare with association rules and to solve cold start and scalable.

CONCLUSION

In this paper, we presented a Recommender System based on complementary and similarity products. Considering the need and taste properties of users, it offers special products. In designed recommender system, two algorithms have been used. CSPAPT offers complementary and similarity products based on previous buying and CSPOPT offers complementary and similarity based on current buying to users. In this paper, we used OWL language to design Electronic Catalogues and calculated the complementary degree based on ontology and offered the need products to users. Recommender Systems can be used in more accurate applications based on complementary and similarity products, such promoting selling strategies and designing selling catalogues to increase in selling system operations.

REFERENCES

1. Yanzan Zhou, Xin Jin and Bamshad Mobasher, 2004. A Recommendation Model Based on Latent Principal Factors in Web Navigation Data, Center for Web Intelligence.
2. Xiaohua Sun, Fansheng Kong and Song Ye, 2005. A Comparison of Several Algorithms for Collaborative Filtering in Startup Stage, Published by the IEEE Computer Society.
3. Mirabedini, J. and M. Khalaji, 2010. Web Usage Mining and Pattern Discovery. Proceedings of WseasConferece, pp: 640-520.
4. Kalles, D., A. Papagelis and C. Zaroliagis, 2003. Algorithmic aspects of web intelligent systems. In: Zhong, N., J. Liu and Y. Yao, (Eds.), Web Intelligence. Springer, Berlin, pp: 323-345.
5. Ardissono, L., *et al*., 2003. Personalized recommendation of tourist attractions for desktop and handset devices. Applied Artificial Intelligence, Special Issue on Artificial Intelligence for Cultural Heritage and Digital Libraries, 17(8-9): 687-714.
6. Huang, Z., H. Chen and D. Zeng, 2004. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. ACM Transactions onInformation Systems, 22(1).
7. Melville, P., R.J. Mooney and R. Nagaragan, 2002. Content-boosted collaborativefiltering for improved recommendations. In Proceedings of the national conference of the American association artificial intelligence (pp: 187-192). Edmonton, Alberta, Canada.
8. Robin Burke, 2002. Hybrid Recommender Systems: Survey and Experiments1, 2002 Kluwer Academic Publishers. Printed in the Netherlands.
9. Knappe, R., 2005. Measures of Semantic Similarity and Relatedness for Use in Ontology-based Information Retrieval, thesis of doctor, Roskilde University, Denmark.
10. Philip Resnik, 1995. Using Information Content to Evaluate Semantic Similarity in a Taxonomy, cmp-lg/9511007 29Nov1995.
11. Moosavi, S.H.., M. Nematbakhsh and H. Khosravi Farsani, 2009. A semantic complement to enhance electronic market, expert system with application, 40(1).