International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 04 | April-2016 www.irjet.net p-ISSN: 2395-0072

# Online shopping recommendation optimization based on users previous search history

**Supriya Jaiswal[1], Aarohi Ajgaonkar[2]**

[1]*Student, M.C.A. Department, Sardar Patel Institute of Technology, Mumbai, India*
[2]*Student, M.C.A. Department, Sardar Patel Institute of Technology, Mumbai, India*

---

**Abstract -**The traditional trading behaviors have been enhanced by the prosperity of e-commerce. Now instead of visiting the shops physically a person feels it more convenient for performing Internet-Shopping. At the same time, consumers experience information overload and look for help in selecting from an unlimited array of products. In order to overcome such types of problem one solution is to develop an online assistance to retrieve product information that really matters for the customers.
Here we basically analyze the users search history based on recommender system. Recommender systems are based on information about a user's past patterns and consumption patterns in general and recommend new items to the user. Also the analysis is done with the help of personalized search whereas this search refers to the search experiences that are tailored specifically to an individual's interests by incorporating information about the individual beyond specific query provided. After performing both the search we apply a Dempster–Shafer theory (DST) rule. This rule derives shared belief between multiple sources and ignores non-shared/conflicting belief through a normalization factor.

***Keywords*** **-**Recommender systems; personalized search; Dempster–Shafer theory (DST) rule;item-based collaborative filtering.

## 1. Introduction

With increasing number of people understanding the true worth and advantages of using the internet, so have the number of E-commerce websites and their customers increased. Online shopping is simply doing normal shopping i.e. physically visiting a store, without the physically visiting part. It is a form of Electronic commerce (E-commerce) that provides the customer the luxury of buying products or services from a seller over the Internet using a web browser. The popularity of E-commerce has increased exponentially over the past few years due to its many advantages such as better prices compared to physical stores, convenience of shopping at own schedule, large variety of goods available online, easy to compare prices of same product among different E-commerce websites, less time consuming and hassle-free, etc.

To help the customers with their online shopping, there are 2 search systems that are made use of – Recommender systems and personalized search systems. Recommender systems study patterns of behaviour to know what a customer will prefer from a collection of things he/she has never experienced based on what other customers who had previously purchased the same product has viewed later. Example can be Amazon's "Customers who bought this also bought" product list. On the other hand, personalized systems display products available for a customer based on his/her previous searched goods. Example is Amazon's "Related to items you've viewed" product list. Styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

## 2. Recommender System

Recommender systems probably are a subclass of information filtering system that desire to predict the user preferences. This system have been used in various applications like movies, music, search queries, social tags etc.This system usually provides a list of recommendations in one of two ways - through collaborative or content-based filtering.
Let us consider an example, if a user is searching on an online shopping website for a new smart phone. So, next time whenever there is a launch of new smart phone with the same or almost same features which user had search for, he will get a recommendation from the same website or from the app regarding this phone. This is how a recommender system works in a short term manner. [7]

Collaborative filtering builds a model from a user's past behaviour (items which are previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users which is then used to predict items (or ratings for items) that the user may have an interest in whereas Content-based filtering is based on description of the item and a profile of the user's

preference in order to recommend additional items with similar properties.

These two approaches are often combined which is known as Hybrid Recommender Systems.

## Collaborative Filtering Model

This is the most widely used approach to design a recommender system which is based on collecting and analysing an enormous amount of information based on users' behaviours, activities or preferences and predicting what users will like which is based on their similarity to other users. It is based on the assumption that people who agreed in the past will also agree in the future, and they will like similar kinds of items as they liked in the past[8].

Let us consider one example, if a user selects his interest for romantic movie and the other user has rated 5 star to a romantic movie on the same movie ticket booking site then the first user will be recommended for that romantic movie which has been rated with 5 star by the other user.

There are two ways to gather the data i.e. to ask for **explicit** ratings from a user, typically on a concrete rating scale (such as rating a movie from one to five stars) or to gather data **implicitly** as the user is in the domain of the system - that is, to log the actions of a user on the site.

1. Item-Based Collaborative Filtering

In collaborative filtering we are making use of model based algorithm i.e. Item-based collaborative filtering which is makes recommendations based on the similarity of items. The problem with user based collaborative filtering was scalability and cold starting problems which is then addressed by Item based. It looks into the set of items the target user has rated and then computes how similar they are to the target item i and then select k most similar items $\{i_1, i_2, \ldots, i_k\}$. At the same time their corresponding similarities $\{s_1, s_2, \ldots, s_k\}$ are also computed. Once the most similar items are found, the prediction is then computed by taking a weighted average of the target user's ratings on these similar items.**[6]**

a) Similarity between the items:

Here the basic idea is to observe all the users who have rated both the items i and j and then to apply a similarity computation technique to determine the similarity $s_{i,j}$.

After finding the similarity between the items we also find the most similar items.[7]

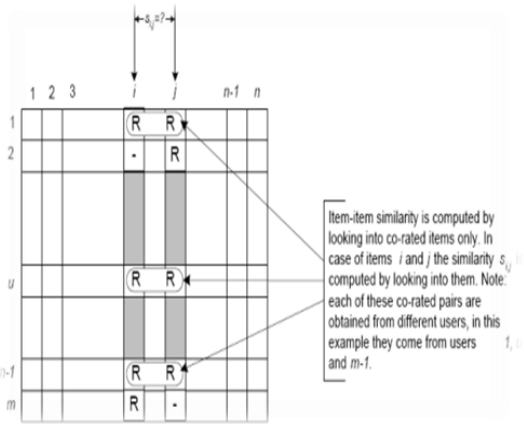


**Fig -1**: Item-Based Similarity

There are a number of different ways to compute the similarity between items. Here we present three such methods. These are cosine-based similarity, correlation-based similarity and adjusted-cosine similarity.

b) Pros of Item based collaborative filtering**:**

- Item-item collaborative filtering provides better prediction than user-user collaborative filtering as the latter dependent on the number of ratings and less number can lead to improper or smaller product results.
- Item-item collaborative filtering i.e. similarity between items is more static compared to that of user-user where the similarity between users is dynamic which repeatedly changes and hence the whole prediction process has to be carried out all over again.
- Easier to implement and scalable as and when new products are added to the online shopping website.
- Rather than matching the user to similar customers, it focuses on the items that customers tend to purchase together.

c) Cons of Item based collaborative filtering:

- Item-item collaborative filtering will not show any useful result in case the user has

not previously visited the website i.e. for first-time customers as there are no previous products searched for that user profile.

3.Personalized Search

Personalization is nothing but providing right information to the right user at the right time. It is done to understand the user and his/her preferences in a much better way in order to help the user in making a fast and accurate choice. This can be done in 2 ways: explicit and implicit. In explicit, the user is asked to provide his choice in the form of ratings or feedback explicitly. Whereas in implicit, the user's behavior over the web is observed and records are maintained without the user being aware. The drawback of explicit is that the user may not provide correct data in a hurry to give feedback and also that the user's preferences may change over time. Thus the feedback may no longer be consistent or accurate.

User browsing history is the most used indicator of user's preferences. If a user visits a webpage, it means that the user has some interest in the contents of that page. The more the user visits a webpage, it more it indicates that the user is interested in the content of the page. Whenever the user searches something over the web, the search results, page ranks, title of the page visited are recorded for future search optimization. Personalization search systems are mostly made use of in the E-commerce websites. With the increase of E-Commerce websites on the internet, it has become tougher for the customers to pick and choose from the wide range of products offered on these websites. Here the customer is left all alone to decide as to what to buy in contrast to in a physical store wherein the information staff or guide is mostly present to help you in choosing the product you have in mind.

To fill the void of a personal assistant in shopping, there is a concept of Personalized Search on the internet. There have been many studies and research done on the various algorithms and their combinations that can be used to build a personalized system. One of them was using a combination of k nearest neighbour (k-NN) algorithm with genetic algorithm (GA). [2]

## Personalized search using KNN and Genetic algorithm

K-NN is a lazy algorithm where the function is approximated locally based on the distance between the nodes. Hence neighbors are found and it is concluded that the behavior of a node will be largely affected and similar to its closest neighbors. Genetic algorithm is a search heuristic that focuses on the process of natural selection. It uses techniques such as inheritance, mutation, selection and crossover. These 2 algorithms were used in the product selection phase. [2]

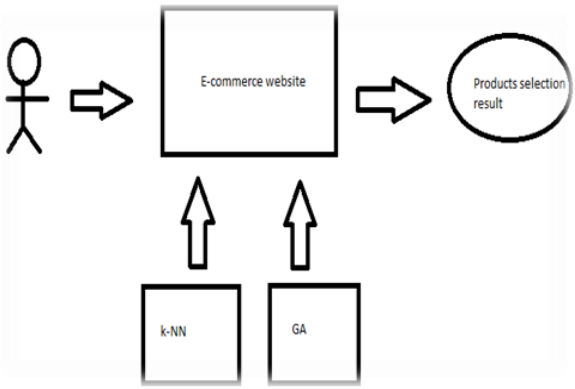


**Fig-2:** Personalized Search using K-NN & GA algorithms.

Here the k-NN and GA work on the data found from the search history of the customer. This history provides a basic overview of the type of products earlier searched by the customer.

The algorithm designed worked as follows: [2]

- Initialize a population of chromosomes consisting of N-dimensional vectors having 12 bits each.
- Evaluate each chromosome according to the kNN classification.
- Create new chromosomes by mating the current chromosomes using mutation and combination techniques.
- Delete old chromosomes that fail to meet the evaluation criteria and develop new chromosomes.
- Evaluate the new chromosomes and add them to the population.
- Stop if stopping criteria is met and return the best chromosome; otherwise go to step 3.

Thus each product is treated as a chromosome having vectors (features or attributes) and only the products which satisfy the evaluation criteria are displayed in the product selection result. Studies have shown this algorithm to be a success in personalized search of products.

## Cases where personalized search can fail to be useful

- The search results of every customer will be different based on the search history or the user profile information of the customer. But sometimes the customer may just want the search results without any personalised search applied to it as similar to when a customer walks into a physical store without any particular product in mind just to randomly checkout the various range of products available in the shop.
- Every click or query searched is recorded to be later used by the personalised search system. If a customer had previously searched and bought a comic of a particular publication, chances are that the personalised search will for a long time show only the comics of the same publication for a long time whenever searched simply for a comic.
- The personalised search will only be helpful if the customer has logged in to the E-commerce website otherwise his/her search will not be recorded accurately. Say another customer using the same machine uses the same E-commerce website, then he/she has to logout the previous user and log in into his own account for the personalised search to be shown. The logging in condition proves disadvantageous in case of a search engine like Google also as people who normally search their queries using it do not tend to always login and hence personalised search is no longer of any use to them.
- Sometimes people search for random stuff online and click on random links when bored. This again is recorded and might affect the results of a personalised search.

### 4.Dempster Rule

Dempster's rule of combination, which combines belief constraintsthat are dictated by independent belief sources, such as in the case of combining preferences. Dempster can be applied on similar type of sources. This rule derives common shared beliefs between multiple sources and ignores non shared/conflicting belief through a normalization factor. We can imagine that several experts indicate their opinion via the attribution of masses of belief $m_1$ and $m_2$ over $\Omega$, or that two successive experiments allow such an attribution[9]. Dempster's rule of combination highlights combined masses of belief in the following way:

$$m_{12}(A) = \frac{\sum_{B \subseteq \Omega,\, C \subseteq \Omega,\, B \cap C = A} m_1(B) \cdot m_2(C)}{1 - \sum_{B \subseteq \Omega,\, C \subseteq \Omega,\, B \cap C = \varnothing} m_1(B) \cdot m_2(C)}$$

The denominator is a coefficient of normalization. In particular, if it is null, it means that there is a total conflict between the sources, and aggregation is then impossible. The use of this rule is thus valid only when the sources are sufficiently in agreement.

### 5.Proposed Model

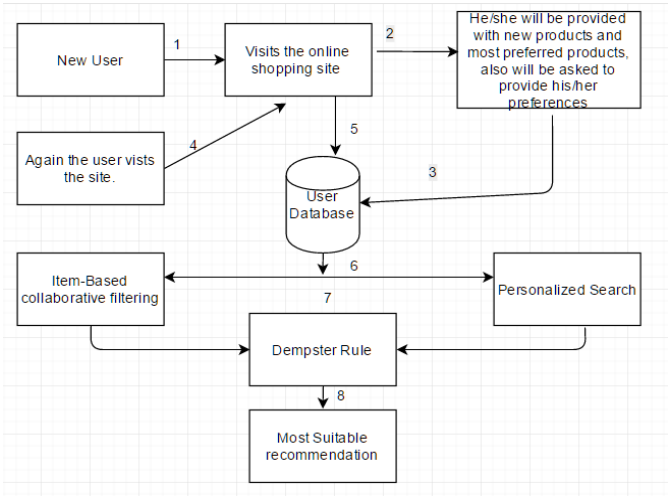


**Fig-3:** Proposed model for most suitable recommendations.

In the proposed model, following steps are considered:

A. New user visits to an online shopping website.
B. As the user is new user he does not have any preferences stored.
C. So he will be shown with the new products and the most preferred products also will be provided with the information of Upcoming products to record his interest.
D. Now this information will be then stored in the database so that it can be used for providing recommendations to user to retrieve product information that really matters for the user and helps in saving the time to search the product from the vast list of products.
E. Now when the user again visits the online shopping site, recommendations will be provided to him.
F. These recommendations are gathered with the help of user based history, after getting the user history

different types of algorithm is applied to it. Two algorithms i.e. Item based collaborative filtering which is a type of recommendation system and personalized search are applied to analyse the data.

G. After getting recommendations as result from both the sources, this recommendation is then provided as an input to dempster rule in order to get the most suitable recommendation.

H. Let us consider an example, User A has shown interest in mobile phones, so suppose recommendation system will provide him two recommendations based on the item similarity i.e. P,Q. Similarly recommendations provided by personalized search can be Q, R. These outputs .i.e. Recommendation System (P, Q) and Personalized Search (Q, R) will be provided as input to dempster rule. According to Dempster rule the probability of recommending product Q is 1.Since it is common to both the algorithm and it is the most suitable recommendation.

## 6.CONCLUSIONS

In this paper, we have indicated the importance of recommendation system and personalized search services in improving the quality of recommendation in electronic-commerce. To realize this we have proposed model which makes use of both the item based collaborative filtering (type of recommendation system) and personalized search which works on user's previous search history. The problem of cold start (This problem is related to recommendations for novel users or new items) is addressed in this model with the help of item based collaborative filtering. The output .i.e. recommendations provided by both of the algorithms are then provided as an input source to dempster rule. Dempster rule works on same type of output format and provides the most suitable results by ignoring the conflicts of the sources. So in this case, dempster rule helps us to combine the recommendation which is common to both the type which becomes the most suitable recommendation thus helping the user to get the best of recommendations which is of his interest and also saves his time. Thus the main agenda of proposed model is that the user should be satisfied with the quality of recommendations & hence the improvement of acceptance rate of online website recommendations.

## ACKNOWLEDMENT

All the authors would like to thank head, department of Master of Computer Applications (M.C.A).We would also like to thank our research guide Prof. Sakina Shaikh for her valuable comments and help towards this research work.

## REFERENCES

[1] JunzhongJi; ZhiqiangSha; ChunnianLiu; Ning Zhong,"Online recommendation based on customer shopping model in e-commerce", conference published in Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on 13-17 Oct. 2003.

[2] Hong-Wei Yang; Zhi-Geng Pan ; Xi-Zhao Wang ; Bing Xu," A personalized products selection assistance based on e-commerce machine learning", conference published in Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on (Volume:4 ) 26-29 Aug. 2004.

[3] Shi ; J. Chen ; Z. Bao;" An application study on collaborative filtering in e-commerce", conference published in Service Systems and Service Management (ICSSSM), 2011 8th International Conference on 25-27 June 2011.

[4] Hsiao ; Yahoo!, Taiwan ; L. J. Li," On visual similarity based interactive product recommendation for online shopping", conference published in Image Processing (ICIP), 2014 IEEE International Conference on 27-30 Oct. 2014.

[5] P.Vashisth ; P.Bedi, " Interest-Based personalized Recommender System", conference published inInformation and Communication Technologies (WICT), 2011 World Congress on 11-14 Dec. 2011

[6] Greg Linden; Brent Smith; Jeremy York," Amazon.com Recommendation Item-to-Item Collaborative Filtering", conference Published in Technology, Education by the IEEE Computer Society Reporter on 2008/11/3.

[7] Daniel Lew;Ben Sowell;Leah E. Steinberg;Amrit S. Tuladhar," Recommender Systems", (Project)

[8] Xiaoyuan Su; Taghi M. Khoshgoftaar;"A survey on Collaborative filtering techniques", conference published in Journal, Advances in Artificial Intelligence Volume 2009, January 2009, and Article No. 4.

[9] Zhang ; B. Deng ; X. Li," Using the Dempster-Shafer theory of evidence to rank documents", conference published in Tsinghua Science and technology(Volume:17,Issue:3) on June 2012.