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. 4, Issue 2, February 2016

Implementation of Recommender System for Customer Personalization

Prema.S, Divakar.S

M.E Student., Dept. of C.S.E., Arunai Engineering College, Thiruvannamalai, India

Asst. Professor, Dept. of C.S.E., Arunai Engineering College, Thiruvannamalai, India

ABSTRACT: Recommender systems be a subclass of information filtering system that seeks to predict the rating (or) performance that a user would give to an item. It have changed the way people find information, products and even other peoples. Among the collection of things people has never practice the study pattern of behavior to know what someone will prefer. Recommender system typically produces a list of recommendation system in one of two ways through collaborative (or) content based filtering. Collaborative filtering method are based on collecting and analyzing a large amount of information on users behaviors, activity or preferences and predicting what users will like based on their similarity for other users. Content-based filtering methods are based on a description of the item and a profile of the user's preferring item. The model needs to calculate the similarity between each pair of item because it suffers from increasing scale of items. To address the issues case based reasoning is proposed to recommend and identify the items that is more suitable for user's buying experience provided that they has selected some items already. Case based reasoning is appropriate to the market basket analysis because the set of item selected together will be adequately compared and modeled. By evaluating the set of selected items, the system can identify the selected item within recommended item in similarity situation and concrete concepts to existing or new user. A case based recommended system will overcome the current limitations cold start and overspecialization problem of content based and collaborative technique by recommending less popular items and generate recommendation to new user. To recommend only new items constraints related to the new items and additional quality characteristics are incorporated and also examine the temporal characteristics of the selected items.

KEYWORDS: Recommendation system,collaborative filtering,content based filtering,case based reasoning.

I. INTRODUCTION

Recommender systems are used in various domains, and become an important part of various applications in order to support both User and Provider for decision making process. It is mainly used in Customer Personalization(meeting the user needs effectively and efficiently)in Market basket data. Most recommendation technique tends to recommend similar user likes and shares similar tastes or similar items those liked by a user in the past. It leads to two major *Limitations:*

(1) Cold Start-shows limited Performance when less Popular or new user appears.
(2) Over Specialization-shows too much additional information about the product.

Market basket domain is characterized by transactions existence of large number of items that consists of Co-Occurring items. Not Simply Predicting whether a single item is liked by a user, the aim is to capture the absence (or)presence of an item within a concept of concrete buying and able to recommend Complementary items. In order to overcome the drawbacks of current recommendation technique there is a need for intelligent recommendation methodologies ,it can generate valuable recommendation based on purchase Patterns and user habits. Case based reasoning is a data mining technique for Problem Solving.CBR main advantages are that it proposes solutions to problem quickly, evaluates the solutions when no algorithmic method is available, helps to stay focused on the major part of the problem and alert to previous experience in solving new and accessible problems. In existing system only provides users, with the products in their stocks and will render the comparison with in their products only. Thereby limiting the users to analyze before buying a product. Existing Service Recommender System suffers from Big data problems like scalability and Time Consumption and thus lack of preciseness.

We Propose a scalable, efficient and Precise service comparison and Recommender system which enables the shoppers to deeply analyze on what product to choose and in which Application, ease and fair with our Gateway. The

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. 4, Issue 2, February 2016

shopper will be provide with clean Indexes of various products with its spec, cost and also Service Ratings which is done in a statistical way. Our System crabs the datas from various web application and loads in its datasets collaboratively and process with batch jobs so as to classify the category and index the data in a distributed and parallel dealing out manner. Shoppers can analyze, Get Recommendations and can pick products and add to cart irrespective of the service provider. Hence our Applications Stands unique as it does not rely on the single Service provider. The Cart can be reviewed at any time and can be processed whenever the shopper wants the product. All the information will be securely and precisely stored in the Users Session. The purchase phase look up for the web services of the products Service provider and can make the online payment with the banks from Service Provider. Once it got over process gets back to our Gateway bringing out the track Id's from product service provider.

II. RELATED WORK

Recommender systems are software techniques and tools for information filtering and retrivel that aims to provide effective and meaningful item recommendations to the active user. The term item refers to the type of entity (information ,service, product,and so on) being recommended; it depends on the application area and on the specific system's objectives.Recommender System regularly cause a set of (top-N) item predictable to be liked by the user or aim to forecast whether a precise item will be of interest.The extensively used recommendation methodologies in commercial applications can be mostly separated into collaborative filtering (CF) and content-based (CB). In CF recommendation techniques,items in the middle of those liked by similar users can be recommended to the new user. A user profile is make of the items that the user has rated highly, thus similarity in user taste are deduce from prior ratings.Even though widely used in marketable applications, collaborative RSs still haveto conquer scalability and cold-start problems that bound their performance.On the other, in CB technique, user profiles are built from the characteristics of the items that a user has rate highly, and the items that he or she hasn't yet tried are compared against them. The items with the highly expected possibility of being liked are then recommended. Because CB techniques rely on more specific information about users and items, they're able to recommend new items. However, they must defeat the recommendations' limited diversity and possible overspecialization.Various hybrid approaches have been proposed to control the strengths of both techniques, defeat their current limitations, and improve their recommendation exactness.

Basic Recommendation Techniques:

- CF (recollection-based, model-based),
- CB filtering (neural networks, probabilistic models, naive Bayes classifier)
- knowledge-based (case-based, confine-based),
- value-based,
- rule-based,
- demographic, and
- other (context-aware, semantically enhanced, crossbreed).

Knowledge-based and especially case-based recommenders enclose emerged as the primary alternative to CF recommenders, intending to overcome their shortcoming while proficiently handling the existing information overload. Case-based recommenders implement a type of CB recommendation that relies on a planned representation of cases, usually as set of well-defined personality with their values. These systems generally recommend items similar to those that the active user has described in his or her request. Rule-based techniques produce item recommendations based on a set of rules extract from a data corpus.Association Rule mining refers to transaction analysis aiming to determine interesting hidden patterns and regular associations among existing items, usually expressed in the form of "if-then" statements.Recently, semantic analysis, latent factors, and probabilistic topic models arising from expected language processing have been effectively applied to information rescue and RSs, especially for tag recommendations .The basic idea is that topics are sets of word given from vocabulary ,and documents are formed as probability distributions over topics .These techniques show privileged accuracy than rule-based options and are able to enhanced handle sparsely problems. Due to the evolution of mobile devices and the use of recommender systems in applications greatly depends on context (location, time, weather, movement state, emotional state, and so on), context-aware recommenders are receiving more awareness. They involve much more than grouping users or items based on their ratings or characteristics—instead, they group users or items coupled with similar context information.

Case Based Reasoning:

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. 4, Issue 2, February 2016

Case-based reasoning (CBR) is a indicative pattern closely related to the human way of reasoning and acting in daily situations when facing novel problems. CBR uses old experiences to solve novel problems, based on the following sentence, known as the CBR assumption:"Similar problems have similar solutions."A Circumstances experienced in the way that it has been captured and erudite is referred to as *past/ previous case* and is stored in the case base. A latest situation asking for a solution forms the description of a *new/target case*. An important part of the CBR methodology is its learning skill, which comes as a natural result of its problem-solving process: the case base is updated every time a new skill is obtained. This knowledge can be reused when needed without implementing the whole process from scrape or to emphasize a methodology that should be avoided in a similar situation. Therefore, case based reasoners are able to progress their problem-solving performance over time. The CBR solving and knowledge process can be described as a cyclical process comprising four processes, known as the CBR cycle, or "the four Rs".

• *retrieve* the most significant cases,
• *reuse* the appreciative provided to the new problem,
• *revise* the result obtained, and
• *retain* the parts of the new solution that are likely to be

used for upcoming purposes.In adding together to the knowledge obtained from previous cases, there's also area-dependent knowledge supporting the CBR process.Cases can be viewed as collected of two parts: the Problem description and the trouble solution.

III. PROPOSED SYSTEM

In this paper four modules are used to describe the system architecture.
[1]. Various web applications Building and Reviewing
[2]. Our Gateway Application and Wed Crawling
[3]. Batch Processing over the TSV Data and other Resources
[4].Picking Products from Recommendations and Purchase

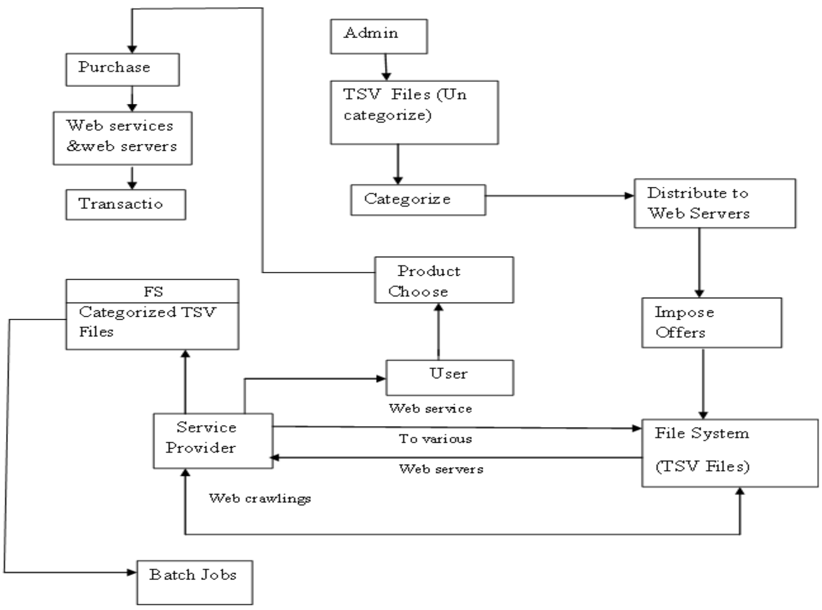

Fig1.system architecture

Various web applications Building and Reviewing:

Sample Web Applications were built so that the users can compare their products with different Service Providers.Similar Datasets were prepared for other Applications too using the meta model that has been crawled

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. 4, Issue 2, February 2016

earlier.Each Datasets were loaded independently in various web Applications. Features and other Specifications have been loaded differently for each Application based on the Service Providers Requirement.

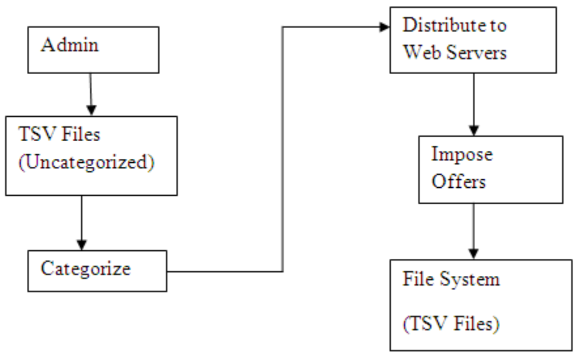

Fig1.a.various web application building and reviewing.

Our Gateway Application and Web Crawling:

Gateway Application is built which gives users with Recommendations and Comparison between the products in the Market.The User can register and can login to view a variety of Products available in Market text a web service method for all service provider.It will Connects to Various Web Application Web Service and can pull all the needed data's to our backend.A huge Amount of data got accumulated now.Web Crawling looks for web service provided by various web applications.

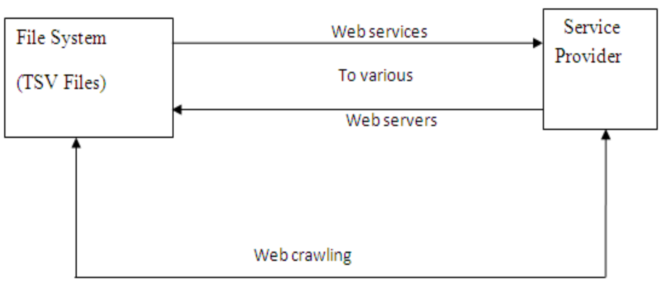

Fig.1.b.our gateway application and web crawling

Batch Processing over the TSV Data and other Resources:

Generally the resources provided by Various Web Servers are in TSV(Tab Separated Values)format and should be Batch Processed before Proceeding.For that use our own API for TSV Manipulation.The TSV files were parsed for data.These data's are used for further processing(recommendation and comparison).

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. 4, Issue 2, February 2016

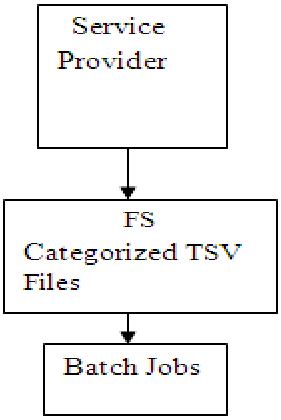

Fig1.c.Batch Processing over the TSV Data and other Resources

Picking Products from Recommendations and Purchase:

The Recommendation were given based on the QOS, Availability, Delivery, Offers , Price and Specifications of the particular product.The User can Pick any product so that our application provides with a most Genuine Recommendation and a set of comparisons.The Users are provided with neat and clean indexes so that they can pick a best provider for a particular product.The picked products were added in cart and can be purchased later.

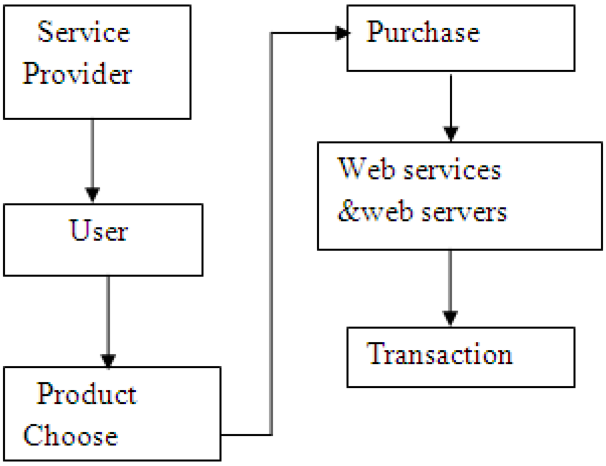

Fig.1.d.Picking Products from Recommendations and Purchase

IV. EXPERIMENTAL RESULTS

To evaluate the implemented recommender, we used a transactional dataset with real market data from a supermarket. The number of offered items was 102,142, with 1,057,076 transactions performed by 17,672 customers. Each transaction is associated with the user who purchased it and the included items. Each item, apart from its name and unique ID, is associated with the various categories that it belongs to (general category, item group, and two item subgroups that, for example, would be, drink, milk, semi skimmed, glass bottle of 1 litre, brand name). This information was used to transform the item representation into a proper depiction, as described above, before generating recommendations. Demographic and subscription information about users can also be found in this dataset. The recommender was tested for different values of parameters that affect its performance, such as the

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. 4, Issue 2, February 2016

number of similar cases retrieved (*k*), the number of recommended items, and the level of detail at which the items are represented. To specify the number of similar cases that would be used, recommendations were generated using 1, 5, 10, 20, and 30 similar cases. The experimental results show that the best option is to use only the most similar case, thus *k* was set equal to 1.We compared the performance of the developed recommender system to three of the widely used techniques, namely, AR, probabilistic topic models, and CF. However, only the first two techniques address exactly the same problem: recommendations of sets of items. CF recommendations focus on the ratings users assign to items and don't take into account joint item selections. But because the CF technique is widely used in commercial applications, its results are also presented. Item descriptions in the transactional database don't contain quality attributes, so we didn't use a CB or a CBR item-based approach. The Apriori algorithm was used to extract ARs from the given transactional database, based on which the recommendations were generated. In addition, we used a topic model recommender (Latent Dirichlet Allocation, or LDA). In this approach, the Offered items are seen as words of a vocabulary, and transactions are treated as documents formed from a combination of topics (item concepts) with some probability distribution. In the CF approach, item selection means the item has a high user rating. We ran various experiments, randomly selecting each time 20 percent of the transactional database for new cases (test set) and the rest as the case base (training set). Using only the most similar case ($k = 1$), Table 1 shows the average results for the recommendation of 5, 7, 9, and 11 items for different representation levels. Because our intention was to evaluate the recommender's ability to identify and recommend the missing items in the transactions, information retrieval metrics (precision, recall, and *f*-measure) were used for the evaluation, with our focus being on the precision value. As you can see, the accuracy of all the methodologies highly improves when using more abstract descriptions. However, as the CBR recommender evaluates the degree of an item's similarity with the items in the target case and not just an item's presence or absence, it outperforms the other recommenders at both representation levels. In contrast, the LDA recommender evaluates similarities among item concepts (topics), while the ARs recommender evaluates only the presence or absence of items within transactions to extract the buying patterns and generate recommendations. Finally, CF recommenders take into account only the presence of items in the user profiles without evaluating the items' co-occurrences within transactions.

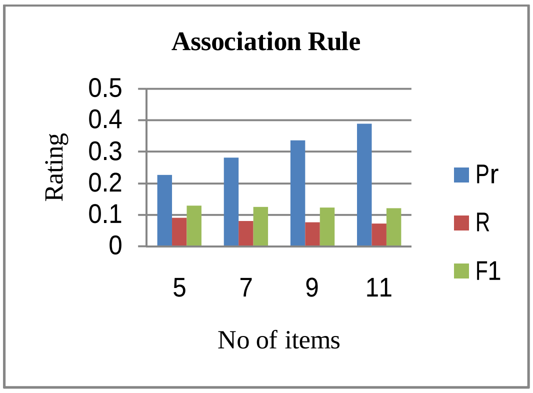

Fig2.a. Association Rule

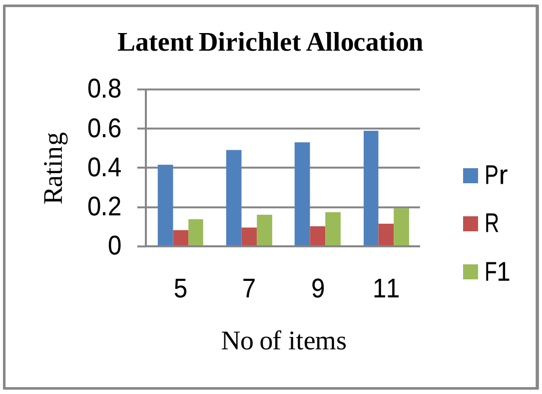

Fig2.b. Latent Dirichlet Allocation

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. 4, Issue 2, February 2016

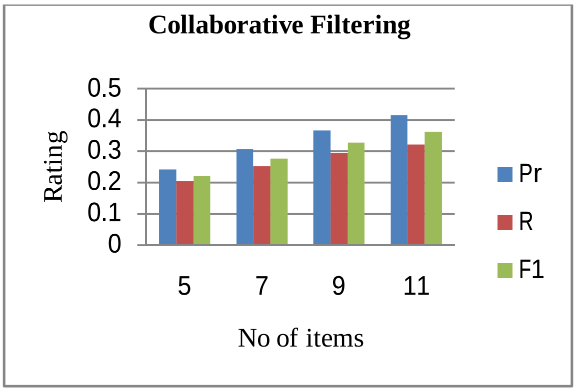

Fig2.c.Collaborative Filtering

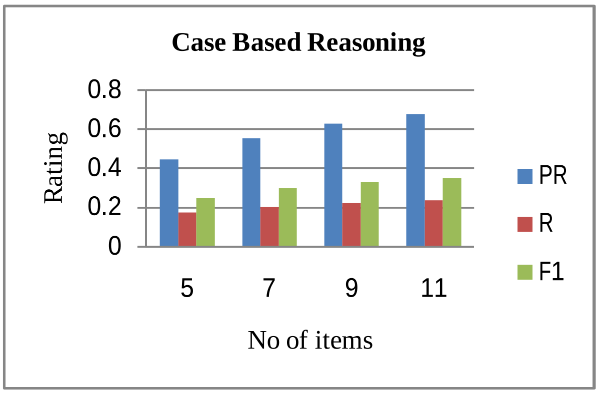

Fig2.d. Case Based Reasoning

V.CONCLUSION AND FUTURE WORK

Recommender systems have become an important part of Numerous commercial applications,enabling customers and providers in their decision-making processes while pursuing their buying and selling strategies. The identification of item selection patterns is thus of high importance. One of our approach's main advantages is its ability to recommend complementary items to the already selected ones. Additionally, it can recommend less popular items and generate recommendations to new users, reducing the cold start and the overspecialization problems of CF and CB techniques. This work could be further extended by incorporating a second processing level into the recommendation methodology. At this level, additional qualitycharacteristics of the items and constraints related to them could be incorporated (for example, to recommend only new items). In addition, it could be applied to other domains where the outcome of a user's experience depends on the total set of items used together. Finally we plan to examine the temporalcharacteristics of the selected items.

REFERENCES

1. F. Ricci, L. Rokach, and B. Shapira,"Introduction to Recommender Systems Handbook," *Recommender Systems Handbook*, F. Ricci et al., eds., Springer, 2011, pp. 1–35.
2. L. Cavique, "A Scalable Algorithm for the Market Basket Analysis," *J. Retailing and Consumer Services*, vol. 14, no. 6, 2007, pp. 400–407.
3. R. Agrawal, T. Imielin´ski, and A. Swami, "Mining Association Rules between Sets of Items in Large Databases," *ACM SIGMOD Record*, vol. 22,1993, pp. 207–216.
4. P. Melville and V. Sindhwani, "Recommender Systems," *Encyclopedia of Machine Learning*, Springer, 2010, pp. 829-838.
5. Z. Huang, D. Zeng, and H. Chen, "A Comparison of Collaborative-FilteringRecommendation Algorithms for E-commerce," *IEEE Intelligent Systems*, vol. 22, no. 5, 2007, pp. 68–78.
6. G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the Stateof- the-Art and Possible Extensions," *IEEE Trans. Knowledge and Data Eng.*, vol. 17, no. 6, 2005, pp. 734–749.
7. F. Lorenzi and F. Ricci, "Case-Based Recommender Systems: A Unifying View," *Intelligent Techniques for Web Personalization*, Springer, 2005, pp. 89–113.
8. M. Steyvers and T. Griffiths, "Probabilistic Topic Models," *Handbook of Latent Semantic Analysis*, vol. 427, 2007, pp. 424–440.
9. J.H. Su et al., "Music Recommendation Using Content and Context Information Mining," *IEEE Intelligent Systems*, vol. 25, no.1, 2010, pp. 16–26.
10. D. Bridge et al., "Case-Based Recommender Systems," *Knowledge Eng.Rev.*, vol. 20, no. 3, 2006, pp. 315–320.
11. K. Christidis, D. Apostolou, and G. Mentzas, "Exploring Customer Preferences with Probabilistic Topic Models," *Proc. European Conf. Machine Learning*, 2010; www.ke.tu-darmstadt.de/events/PL-10/papers/3-Christidis.pdf.