IOSR Journal of Computer Science (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727
PP 54-61
www.iosrjournals.org

Semantic Analysis of Context Aware Recommendation techniques

Mohammad Daoud[1], S.K Naqvi[2], Alok Nikhil Jha[3]

[1]*Department of Computer Science and Engineering&Centre for Education Technology, Indian Institute of Technology, Delhi, New Delhi 110016, India*
[2]*FTK-Centre for Information Technology, Jamia Millia Islamia, New Delhi 110025, India*
[3]*Innovation & IT, Cluster Innovation Center, University of Delhi, Delhi 110007, India*

ABSTRACT: *The paper entails the existential importance of recommendation engines & showcases the dependency of electronic business on feedback mechanism & recommendation systems evaluation methodologies. The paper also presents the semantic evaluation & analysis of the available algorithms for recommendation engines, strategies & technologies used in backend engines on web server or as application server within a domain. Finally the paper discusses our proposed model of recommendation with the aspects for recommending any object on the digital domain with provisions of subjective evaluation & measurable semantic relationship considering the sentiments of the users, scaled to a self-learning model.*

Keywords: *Recommendation Parameters, Recommendation System, Semantic similarity, Sentiment Analysis, User generated feedback*

I. INTRODUCTION

With increased penetration of Internet and Web Technologies, life has become much easier. The judgemental attitude and professed decision has started coming from every individual who are not too much into technologies but are enabled to take technological decisions in terms of buying a gadget. With more usage of devices, tools, gadgets & emergence of ecommerce, more data are uploaded and transacted on Web every day.

A huge amount of Data on individual topics are uploaded every day and searching for keywords on a search engine results large chunk of Data each day unlike 10 years back.

The results showcasing pattern has been constant for years but still the changes in the result procurement, when a keyword is searched, is not so through correct always. The results may differ when a same user is logged into portal and when access it anonymously. There by a strong need of system exists that would enable the user in selecting right product, services and understand their correct requirements and make a worthy decision in buying the product.

The Recommendation Systems have been conceptualized and are providing the Suggestions to End Users, customers, competitors and companies itself. Right after the conceptualization of Recommendation systems, they are available in different domains for different purposes in analysing the data available with user's experience and suggesting the users to make a professed decision which is based on personal experience of other users.

With the availability of such system, when a user plans to purchase any item online or offline, they undergo a process of surveying about the product, stores, pros, cons, offers and these information are available by companies but most of the relevant information i.e. meaningful for user comes up from recommendations& feedback, reviews of the people who are using it or are expert in it or else carry some knowledge about it apart from feedback of the company or seller.

The different Recommendation systems uses Suggestion Algorithms designed as per the focussed product domains. We are working on a Dynamic Recommendation system that has its unique Suggestion Algorithms which has a collaborative learning factor and modification of algorithms with the incoming feedback, inputs and data from users or/and customers. Our recommendation system would basically parse feedbacks and reviews into two segments i.e. Feature & Opinions and implement it with integrated learning plans with categorization of sentiments & evaluate correct perception with the language & words used to stratify the facts.

II. LITERATURE SURVEY

Recommender systems have been implemented in a broad range of domains, including news [1], movies [2], music [3], books [4], laptops [5] and so forth. Nowadays, recommender systems exist in a variety of areas in our daily life.

For example, Netflix (netflix.com) provides movie recommendation services; Last.fm (last.fm) and Pandora (pandora.com) recommend music; and Amazon (amazon.com) recommends miscellaneous retail products such as books, electronics. Since recommendation was considered an independent research area in the mid-1990s, a number of recommendation approaches have been proposed. They can be mainly classified into two types [6]. Collaborative filtering (CF) is one of the most successful recommendation technologies. The basic idea behind this method is that it gathers the opinions of others who share similar interests with the target user. Another principal recommendation technique is called content-based filtering, which is originally derived from information retrieval technologies. Content-based approaches predict which items the active user would have rated highly, by matching his/her preference profile with the attributes of the items to be recommended. Since both technologies depend on user ratings, they cannot work well when the number of the ratings for a user or an item is inadequate [6]. For example, when a new user enters a system and gives no or few ratings to items, it is difficult for the system to accurately predict this user's preferences and provide reliable recommendations. It is referred to as the “new user" problem. New items have the same issue, known as the “new item" problem. To address this problem, other modelling resources have been incorporated into recommender systems, for example, domain knowledge [7, 8], tags [9] and human demographics [10].

Demographics techniques to handle the new customer problem. Demographic techniques form “user-to-user” correlations like collaborative ones, but use different data. The benefit of a demographic approach is that it may not require a history of user ratings of the type needed by collaborative and content-based techniques [10]. After getting the feedback from the customer Researcher analysis that system recommends all those products they have high rating or match against the customer preferences without knowing the knowledge about customer such as time, place and company of customer etc. For example the system recommends a vacation package to customer [11]. System recommends those places they have high rating or match against the customer preferences without knowing the time period. Suppose the customers have a vacation plan in winter then the system recommends also those places where the customer goes to prefer in summer or recommends that place where the customer goes to prefer in winter. Knowledge base system handle this problem the system should have the knowledge about the object and customer needs with time, place and customer needs. Knowledge-based approaches are distinguished in that they have functional knowledge: they have knowledge about how a particular item meets a particular customer need, and can therefore reason about the relationship between a need and a possible recommendation [10]. E-commerce websites can predict a customer's future purchasing behaviours through the information collected from a customer's past purchasing behaviors and demographic data of that Customer. Those customers with similar purchasing habits and interests are grouped together in order to recommend products with surprises to target customers. The candidate recommendations in the two major parts are presented to target customers through a recommendation System. With the subsequent customers' feedback responses, the items of recommendations in the system are adjusted [12]. To predict the customer future purchase the system recommends all those products that's matching customers' needs but that community unable to completely satisfying the customers. Then Marginal utility concept added. Marginal utility is economic concept because economists and marketing research use it to determine how much of an item a customer will purchase. According to the Law of Diminishing Marginal Utility, many products have the decreasing marginal utility with the increase of purchase count, such as cell phones, computers, and so on. Users are not likely to purchase the same or similar product again in a short time if they already purchased it before. On the other hand, some products, such as pet food, baby diapers, would be purchased again and again. Then marginal utility will be increase [13]. Using this concept the customers become are more satisfied. Another technique is Hybrid methods in which at least two of the existing techniques are used to gain better performance with fewer of the drawbacks of any individual one. The aim is to take benefits of all techniques and obtain more relevant suggestions and improve recommendation performance. The most popular approach of Hybrid methods is content based systems with collaborative filtering systems.

Further researcher aim to resolve the recommendation problem by using the context based cascaded ratings having semantic similarity in online environments when customer rating information is not available or keeps on adding up in the system. In this Researcher work on the investigation of customer reviews that broadly appear in website, they hence propose a new recommender algorithm by fusing a self-supervised emotion-integrated sentiment classification approach, by which the missing User-Item Rating Matrix can be substituted by the virtual ratings which are predicted by decomposing customer reviews as given to the items[14]. Using the concepts we move forward and the user generated content called "review" plays an important role both for the buyer as well as for the seller. Further in our work we are mining these reviews and extracting the features and opinions with modifier (if exit).using these feature and opinion we conceptualize an architecture of analysis of semantics behind the similarity of context & feedback.

III. ISSUES IN EXISTING TECHNOLOGY & IMPLEMENTATION

The most of users & competitors are relying on recommendation system but the existing systems implemented are following the techniques such as collaborative method, content base method, knowledge base systems, demographics, utility, rule base and hybrid methods, which are now somehow traditional techniques or not so effective as the set of items, set of users & set of reviews all are increasing exponentially. The above stated methods and recommendation systems based on these algorithms were more focussed on items having specific set of attributes e.g. cost, availability, branding, quality, servicing (if applicable), count of users using it. The trends were lesser in case of products which are rarely purchased or not carrying that credibility in market or among customers base. The mind set with a low rating of product(s) in terms of recommendation systems assigning the values on basis of feedback & reviews carry a low status which definitely creates an impact on the Company providing the product or solution and users who check the macro level facts. Users focus goes on ratings provided by recommendation engines designed by the service providing company or magazines, other media which carry out people survey where in Sample size of data is ignored while considering those surveys.

Another issue in the systems are directly proportional to language and words used for specifying the features & personal opinions of the product. These engines once designed and are on field never checks the dynamism of linguistics of region. For example, consider a product’s specific case with multiple inputs from different users

- “ item is not too bad”
- “item is not too good”
- “item is worth trying”

All the statements provides a different inference and different recommendation system would perceive them differently. The changing linguistics culture worldwide is not catered in any of the recommendation algorithms and hence lags extension in the existing systems.

The existing recommendation system lacks the capability of creating the difference on respective systems specifications meaning a review on specific attribute of system is utilized to measure the overall system or item’s rating. To quote an arbitrary example not having an ergonomic mouse cannot account for performance of computer or not having metallic buttons cannot quote of quality of a shirt. These kinds of feedbacks are accountable in existing systems while rating or comparing the complete product(s).

IV. PROPOSED SYSTEM

The accountability of product review mechanism dependency should be catered independently for all of its attributes. Our recommendation engine focuses on the product’s features and recommends & condemns it depending upon its feature and requirements of users fairly considering the reviews in the system. We are dividing the products specifications reviews into feature set and opinion set. The feature set would be features/attributes of any product or system and Opinions set would be people’s comments/reviews. The feature and Opinion would be mapped together for different products on basis of reviews from different users. Our system would use the existing techniques of mining the Feature Set & Opinion Set from the available research work [15].

We make the feature and opinion set “F_{set}” and “O_{set}” respectively & the intra relation by matrix representation with the help of Word sense based Clustering which is shown in figure1.

$$
\begin{array}{cc}
 & \begin{array}{ccccccccc} F1 & F2 & F3 & F4 & F5 & F6 & F7 & \cdots & Fn \end{array} \\
\begin{array}{c} F1 \\ F2 \\ F3 \\ F4 \\ F5 \\ F6 \\ f7 \\ \vdots \\ fn \end{array} & \left[\begin{array}{ccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & . \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & . \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & . \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & . \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & . \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & . \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & . \\
 & & & & & & & & \\
. & . & . & . & . & . & . & \cdots & 0
\end{array}\right]_{N\times N}
\end{array}
$$

Fig 1

The other step involves identification of semantic similarity in text clustering of opinion set i.e. words may have a sense & display a canonical relationship (F1 and F7 have same sense) as in figure1 or transitivity factor i.e. F2, F4 have same sense and F2, F6 have same sense then transitivity occur in F4 and F6 in feature set as well as opinion set.

We clustered word together those have same sense in both set F_{set} and O_{set} .

Example: 1 - "laptop keyboard excellent ".

2- "Keyboard is too good"

3- "Laptop keypad bad"

4 – "Keypad working better"

5 – "Buttons sense very cool"

We got above reviews from users for a specific laptop product. At first stage we need to cluster the words. Some of these words give same sense but some have different but still the statements are focused on attribute and same attributes are referred by different name.

1- Statements have different term for feature and opinion and we have need to find the sense of each feature as well as opinion. While clustering the sense we may face some transitivity occurring.
2- As in review no. 3 "laptop keypad bad". This user has different opinion of same feature & hence in this situation we cannot neglect the user view.
3- Generally we recommend users that attribute is good (as 80% user have +ve feedback and 20% have –ve feedback on this feature) with the sample size.
4- We'd identify intra relationship in Feature set as well as Opinion set as a matrix. And to find the semantic similarity between the feature as well as opinion we use the Adjacency list representation. The methodology maintains space proportionality to |E| + |V| (so proportional to the size of the graph) where E represents Edges & V represents set of Vertices.

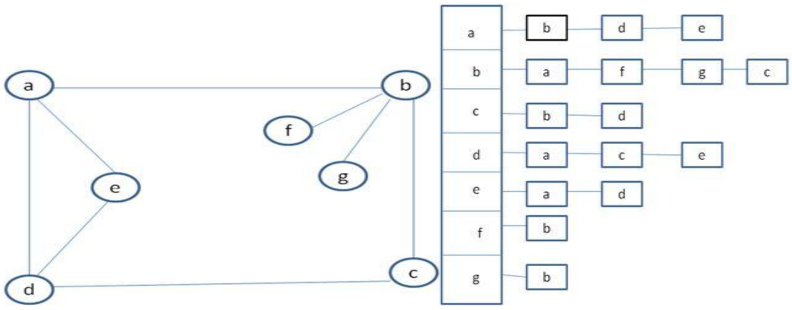
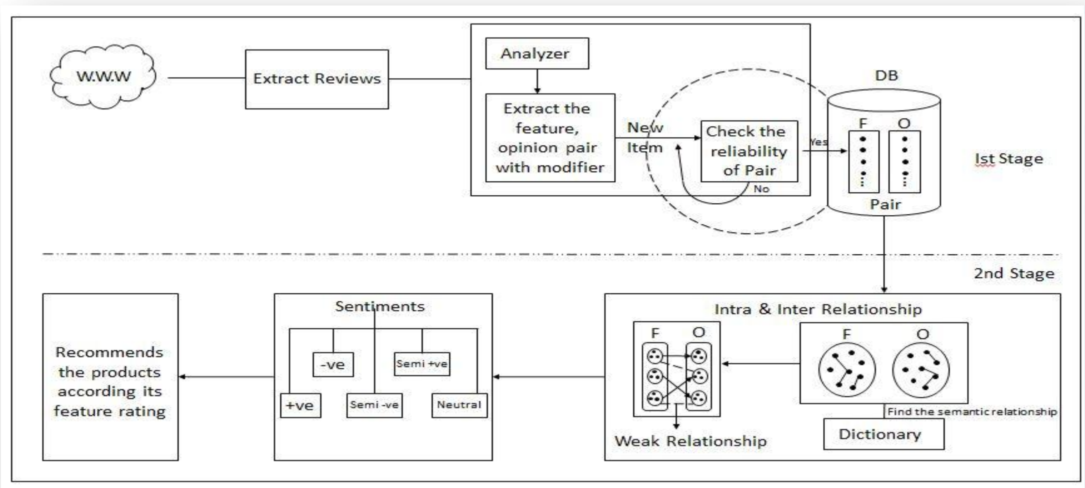


Figure 2

Once we have identified the Feature Set & Opinion Set in our architecture, we setup a Feedback mechanism wherein the new values from the source data are fed back with new values coming from same or distinct users, so that we have a diversified source of learning mechanism integrated in the system in place with the flexibility to manage the language gaps in explaining & illustrating the feedback, thereby, constituting a better set of Features & Opinions.

In the Stage two of our architecture, we identify the inter relationship and intra relationship & correlate with the semantic similarity among them with a synchronized word sense system to understand the context of reviews which are in form of Features & Opinions. The features & opinions are segregated to evaluate the sentiments of the users in context of specified component of the service or product and accordingly the feedback would be converted to an objective feature from subjective inputs.



The whole feature, opinions are used to calculate the Sentiments of the users towards the product and the sentiments are basically categorized into 5 subparts which are cited below:

I. Positive
II. Negative
III. Semi-Positive
IV. Semi-Negative
V. Neutral

Positive sentiment implies that calculated implication is completely positive & gives a positive sense of feedback and semi-positive state is dependent on inputs received from the users recursively that is with the cascaded feedback mechanism initiated at segmentation of Feature & Opinion Set in Stage-I. Negative & Semi-Negative goes in the same way. The sentiment analysis provides overall rating on the basis of an algorithm which keeps updating the self with the inputs from stage-1 and update in word sense dictionary synchronizing the information in real time. The system learns the mechanism of analysis and delivers on the fly with the correct data set and thus provides a full proof system of implementation.

V. CONCLUSION

The architecture framework implies a direct modification on existing techniques as the system has dependency on cascaded network of inputs in terms of reviews & feedbacks from end users which makes it a better model of analyzing the semantic relationship in real time. The sentiment analysis follows context awareness structure and would be catered in real time and we are designing the algorithm for better results.

REFERENCES

[1] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative ltering of netnews, *In Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW'94), pages 175{186, New York, NY, USA, 1994. ACM.*

[2] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, An algorithmic framework for performing collaborative ltering, *In Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 230{237, New York, NY, USA, 1999. ACM.*

[3] U. Shardanand and P. Maes,Social Information Filtering Algorithms for Automating Word of Mouth, *In Proceedings of the SIGCHI Conference on Human factors in Computing Systems (CHI'95), pages 210{217, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.*

[4] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, Improving recommendation lists through topic diversication, *ProceedingWWW '05 Proceedings of the 14th international conference on World Wide Web*

Pages 22-32.

[5] L. Chen and P. Pu, Hybrid critiquing-based recommender systems, *In Proceedings of the 12th International Conference on Intelligent User Interfaces (IUI'07), pages 22{31, New York, NY, USA, 2007. ACM.*

[6] G. Adomavicius and A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE Transactions on Knowledge and Data Engineering, 17(6):734{749,June 2005.*

[7] R. D. Burke, K. J. Hammond, and B. C. Young, Knowledge-based navigation of complex information *space,. In Proceedings of the 13th National Conference on Articial Intelligence (AAAI'96), pages 462{468. AAAI Press, 1996.*

[8] P. Pu and L. Chen, User-involved preference elicitation for product search and recommender systems, *AI Magazine, 29(4):93{103, 200, Vol 29, No 4.*

[9] R. J□aschke, L. Marinho, A. Hotho, S.-T. Lars, and S. Gerd, Tag recommendations in folksonomies. *In Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, pages 506{514, Berlin, Heidelberg, 2007. Springer-Verlag.*

[10] R. Burke, Hybrid recommender systems: Survey and experiments, *User Modeling and User-Adapted Interaction archive Volume 12 Issue 4, November 2002 Pages 331 – 370*

[11] R. Burke, Knowledge-based recommender systems, *To Appear in the Encyclopedia of Library and Information Science 2000.*

[12] S-S Weng , M-Ju Liu , *Personalized Product Recommendation in E-Commerce , Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04).*

[13] J. Wang, Y. Zhang , Utilizing Marginal Net Utility for Recommendation in E-commerce, *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ISBN: 978-1-4503-0757-4.*

[14] W. Zhang, G. Ding, Li Chen, Chunping Li , Augmenting Chinese Online Video Recommendations by Using Virtual Ratings Predicted by Review Sentiment classification, *Proceeding ICDMW '10 Proceedings of the 2010 IEEE International Conference on Data Mining Workshops Pages 1143-1150.*

[15] A. Kamal, M. Abulaish, T. Anwar, *Mining Feature-Opinion Pairs and Their Reliability Scores from Web Opinion Sources, Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics Article No. 15*