International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-8)

Research Article

August 2015

Online Product Prediction by Utilizing Social Network Features

Monika Pal, Srikant Lad

Abstract—

Online Social Rating Networks such as Epinions and Flixter, allow users to form several implicit social networks, through their daily interactions like co-commenting on the same products, or similarly co-rating products. This paper work on the product prediction of the user base on the social network, user product rating. Here new concept of jaccard coefficient is involve in the work, where different features of the user from social graph is use for prediction. Results shows that proposed work has high precision and recall value on different dataset size as compare to previous approach.

Keywords— Data relation, Jaccard Coefficient, Product prediction, social network, social features, .//minimum 5 keyword

I. INTRODUCTION

As digital world include different facility for the ease of human life, one of its boon is internet. This internet has reduced hundred of works, in terms of managing things, providing information, etc. So number of internet users are increasing day by day, many of different kind of services are developed through which people maintain their social relation buy products, sell old products, etc. Utilization of this traffic is done by researcher by analyzing user reviews and comments for product purchase prediction [2,5,9].

Many of sites like epinion, flixter, etc. are maintaining users ratings for variety of products. Users co-comment on same number of products share its experience which can increase or decrease products rate. On these sites users are from different social networks but communicate each other on the basis of that product relation. Opinion of one lead to one long continuous talk help other for purchasing.

Social networking sites like facebook, twitter, etc. provides good platform to maintain relation among people. On these sites people put their personal life. So utilization of both kind of sites are done in [8, 10]

It is known that people buy those products that are suggested by some other person who has trust on it. As advertisement for particular community increase product selling if target community is correct. So product prediction work will identify those community for advertising. As chance of purchasing that kind of product is high.

Chance of product sale is highly based on product feature. Its sale increases by making proper advertisement. Product prediction accuracy increase by user reviews, while combination of social network in product prediction will also increase accuracy as done in [1, 8].

Problem Identification: This work focus on product recommendation system, where user recommend product by rating products. Utilization of this single recommendation has low efficiency. So some of researcher has included social relation trust value, where user-user trust is present. But direct specifying the trust is difficult as it vary with time. So proper trust calculation is required. In this work use of fuzzy interval value is done for calculating trust from current user social features value with other. So combination of rating and user trust is done in this work for increasing the efficiency of prediction.

II. RELATED WORK

Product recommendation system is done by two main techniques first is content based and other is community-based. Content-based technique is popular as compare to community base although some interest has focused on collaborative filtering [12]. In content base recommendation users own preferences is specify, so retriving decision from this is easy and perfect [19].

In [12] frequency vector of users are maintain by hashtag and entities specify in tweets. Similarly one more vector having URL links of twitter is maintained in separate vector. So Users are then recommended URLs whose vector is most similar to theirs.

In [14] BOW (Bag of word) is maintain for the common words use in URLs and twitter terms specify by the users. Here social network among users is not consider.

In [18] friendship base random walks is done for collecting data from tweets by users. Here similar kind of tweets are utilize for the user-user recommendation.

In [16] similar approach of random walk is adopt but technique use for recommendation is base on collaborative filtering instead of content-based techniques.

In [17] recommendation is done by combining user influence and user recommendation. Based on this probabilistic matrix model prediction is done that is both similar in nature. Here it is obtain that Community-based systems has better accuracy as compare to content-based and collaborative filtering methods.
Social Networks and Purchase Behavior is analyzed by some researcher and has investigated the broader topic of how social network influences users in their purchases. In [13] empirically demonstrate that a user's friends exercise \peer pressure": if friends widely adopt a product, the user is more likely to buy it.
In. [15] study the trading dynamics on the e-commerce social network Taobao. They show that buyers are more likely to purchase from sellers that friends in their network have already bought from (information passing). They prove that when a buyer has to decide from which seller to buy a product, the social network has a bigger inuence on the decision than the sellers' ratings and the price of the product.

Basic Notation

Whole work focus on social feature base trust development and utilize this trust for rated item prediction. As purpose of social network is varying from site to site so number and type of feature also vary. In this work facebook social network is consider and its feature set is consider for trust calculation. Here many events are as comment, like, tag, unlike, write on wall, etc. Here each type of event is act as feature. This

III. PROPOSED WORK

In this work product is predict, with the use of different relation such as user product relation user user relation. Base on these relation a new combination of features is use for the prediction of product that will be purchase by user. So fig. 2 represent the steps of proposed work.
User-User Datset: In this dataset user user feature relation is present. This can be understand as user U1 has some relation with U2 in terms of {Like, comment, share image, shar video, message, share comment, friend request, same group, common friends, video chat, text chat, etc.}, then number of time these activity done by the user is count in the dataset for U2 by U1 is store.
Pre-Processing
As dataset contain number of feature between user so conversion of dataset as per working environment is done in this step here dataset is arrange into matrix form where first two column represent user is while rest of column represent the feature count value present in dataset. If zero present in the column then it shows that that feature is not use by the specify user ids in first two column.
UUD ← Pre_processing(UUD)
Jaccard Coefficient.
The value of the features are is in integer form and differ person to person, so the Jaccard Coefficientare generate from the pre-processed dataset:

$$\textbf{Jaccard-coefficient}(x,y) = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|}$$

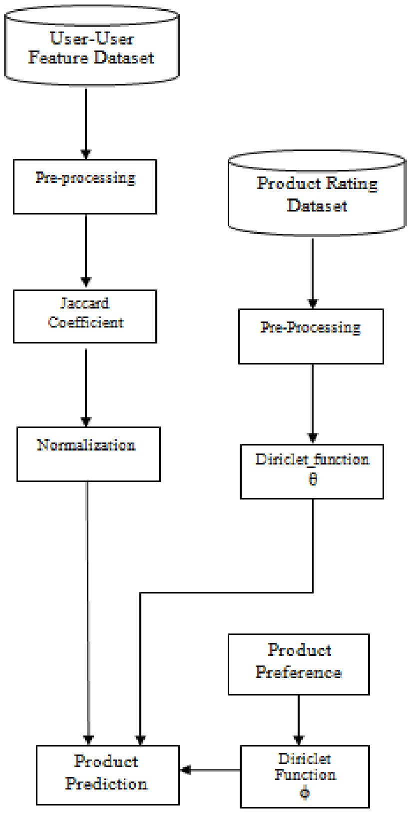

Fig. 3 Block diagram of Proposed Work

Conceptually, it defines the probability that a common neigh
bor of a pair of user x and y would be selected if the selection is made randomly from the union of the neighbor-sets of x and y . So, for high number of common neighbors, the score would be higher. This can be understand as let X∩Y has like feature value 5, while XUY has like value 50 then Jaccard Coefficient value is 0.1.
JC[]←Jaccard_Coefficient(UUD)
Normalization:
In this step matrix obtain after the jaccard coefficient need normalization. As different feature has different priority so to put all feature in same scale normalization is required.
So if W is the weight matrix, then following step will do normalization.
```
Loop 1:n
JC[n] ← JC[n]*W
EndLoop
```
Product Rating Datset
In this dataset product rating feature is present. This can be understand as user U1 has either use or have knowledge or its review for any product id P1then rate it on the basis of his thought such as {best, very good, better, good, ok}.
Pre-Processing
As dataset contain number of rating between user and product so conversion of dataset as per working environment is done in this step here dataset is arrange into matrix form where first column represent user-id second represent product-id while third us for rate. For giving rate instead of giving any text rate values are provide for each class. If zero present in the column then it shows that that product is not use by the specify user ids.
UPD ← Pre_processing(UPD)
Latent Dirchlet Algorithm
Here with the help of this function dirrchlet will give a value as a relation between the user and user or item which is base on the UPD rating dataset.
θ← LDA(UPD)
In the similar fashion each product has its own product preference, so by the use of LDA one more relation is introduce.
Φ← LDA(Product_preference)
Product Prediction
This is the final step here user product prediction is done on the basis of social graph (user-user dataset), user-item relation, item preference.
In this step each user Xn who is friend of user Yj where j=1,2….t where t is number of X friends.
```
Loop 1: n
        Loop 1:j
                P[j] ← θj*ϕ*JCn
        EndLoop
EndLoop
```
Now this P has j number of entries. So the maximum value index in P will be the final product id.
Proposed algorithm:
Input: UUD, UPD, Product_Preference
Output: Product_prediction

1. UUD ←Pre-Processing(UUD)
2. JC[]←Jaccard_Coefficient(UUD)
3. Loop 1:n
4. JC[n] ← JC[n]*W
5. EndLoop
6. UPD ← Pre_processing(UPD)
7. θ← LDA(UPD)
8. Φ← LDA(Product_preference)
9. Loop 1: n
10. Loop 1:j
11. P[j] ← θj*ϕ*JCn
12. If P[j]>T
13. [x y]←Find_Friend(UPD[j], n) // Find friend of user n for product j
14. P[j] = P[j]*(x*C1 + y*C2) // x is Number of j product user who are n friend and y is other user of product j
15. EndIf
16. EndLoop
17. EndLoop

IV. EXPERIMENT AND RESULT

Experimental Setup

This section presents the experimental evaluation of the proposed work. All algorithms and utility measures were implemented using the MATLAB tool. The tests were performed on an 2.27 GHz Intel Core i3 machine, equipped with 4 GB of RAM, and running under Windows 7 Professional.

Dataset

The Epinions dataset contains

- 49,290 users who rated a total of
- 139,738 different items at least once, writing
- 664,824 reviews.
- 487,181 issued trust statements.

Users and Items are represented by anonimized numeric identifiers.
The dataset consists of 2 files:first file contains the ratings given by users to items, second file contains the trust statements issued by users.

Evaluation Parameter

To test outcomes of the work following are the evaluation parameter such as Precision, Recall and F-score.
Precision = TP / (TP+ FP)
Recall = TP / (TP + TN)
F-score = 2 * Precision * Recall / (Precision + Recall)
Where TP : True Positive
TN : True Negative
FP: False Positive
Area Under the Curve (AUC):
With the help of precision and recall value AUC value is calculated that is term as the Area under the precision, recall curve.

Results

Results are compare with the previous work in [1] whicg=h is term as previous work in this paper.

Table. 1. Comparison results of Previous work with proposed work for 300 user and 1000 product.

Values	**Previous [1]**	**Proposed**
Precision	0.1176	0.4231
Recall	0.0132	0.0679
F-Measure	0.0237	0.1170

Table 2 Comparison results of Previous work with proposed work for 500 user and 1000 product.

Values	**Previous [1]**	**Proposed**
Precision	0.2069	0.3429
Recall	0.0251	0.0482
F-Measure	0.0448	0.0845

It has been observed by table 1 & table 2 that by product prediction of proposed work is better as compare to previous one, as precision value is higher. It is observed that as the size of the datset increases then number of user and there chance of generating product prediction get less. This due to the confusion or the randomness of user.

.

Table. 3. AUC of the previous and proposed.

Area Under Curve		
Number of Users	**Previous [1]**	**Proposed**
300	0.012	0.0144
600	0.0026	0.0083

It has been observed by table 3 that by product prediction of proposed work is better as compare to previous one, as AUC of different evaluation parameter from table 3. It is observed that as the size of the datset increases or decrease values are always above from previous work.

V. CONCLUSION

As the increase of social media, number of internet user has increase in large number. This enhance e-shoping, so researchers get new field for mining that is product prediction. Web item prediction has been widely used to reduce the user confusion problem. This paper has focus on product prediction where new combination of Jaccard base social network utilization is done with probabilistic function LDA. Results shows that with the increase in features for Jaccard coefficient prediction accuracy has increase. Although research in this field is just a start, it is required to develop an adaptive algorithm as per social network.

REFERENCES

[1] Freddy Chong Tat Chua, Hady W. Lauw, and Ee-Peng Lim. "Generative Models for Item Adoptions Using Social Correlation". IEEE transaction on knowledge and data engineering, vil. 25, no,. 9, September 2013.

[2] Katz, Leo. (1953) A new status index derived from sociometric analysis. Psychometrika, 18(1):39-43.

[3] J. Herlocker, J. Konstan, A. Borchers, and J. Riedl. An algorithmic framework for performing collaborative filtering. In *Proc. ACM SIGIR Conf.*, pages 230–237, 1999.

[4] Liben-Nowell, David, and Kleinberg, Jon. (2007). The Link Prediction Problem for Social Networks . Journal of the American Society for Information Science and Technology, 58(7):1019-1031.

[5] J. Herlocker, J. Konstan, L. Terveen, and J. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. on Information Systems*, 22(1):5–53, 2004.

[6] M. Jamali and M. Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In *Proc. 4th ACM RecSys Conf.*, pages 135–142, 2010.

[7] L. Kartuz. A new index derived from social analysis. *Psychometrika*, 18(1):39–43, 1953.

[8] H. Li, S. Bhowmick, and A. Sun. Affrank: Affinity-driven ranking of products in online social rating networks. Journal of the American Society for Information Science and Technology 2012.

[9] D. Liben-Nowell and J. Kleinberg. The link prediction problem for social networks. In Proc. 12th CIKM Conf., 2003.

[10] P. Massa and P. Avesani. Trust-aware collaborative filtering for recommender systems. In Proc. Federated Int. Conf. on The Move to Meaningful Internet:CoopIS, DOA, ODBASE, pages 492–508, 2004.

[11] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proc. WWW Conf.*, pages 285–295, 2001.

[12] [1] F. Abel, Q. Gao, G.-J. Houben, and K. Tao. Analyzing temporal dynamics in twitter pro_les for personalized recommendations in the social web. In WebSci'11 Conference Proceedings, 2011.

[13] R. Bhatt, V. Chaoji, and R. Parekh. Predicting product adoption in large-scale social networks. In CIKM'10 Conference Proceedings, 2010.

[14] J. Chen, R. Nairn, L. Nelson, M. Bernstein, and E. Chi. Short and tweet: experiments on recommending content from information streams. In CHI'10 Conference Proceedings, 2010.

[15] [11] S. Guo, M. Wang, and J. Leskovec. The role of social networks in online shopping: information passing, price of trust, and consumer choice. In ACM EC'11Conference Proceeding, pages 157{166, 2011.

[16] [15] M. Jamali and M. Ester. Trustwalker: a random walk model for combining trust-based and item-based recommendation. In KDD'09 Conference Proceedings, pages 397{406, 2009.

[17] [18] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang. Social contextual recommendation. In KDD'12 Conference Proceedings, pages 45{54, 2012.

[18] [34] R. Yan, M. Lapata, and X. Li. Tweet recommendation with graph co-ranking. In ACL'12 Conference Proceedings, volume 1, 2012.

[19] [16] D. M. Romero, W. Galuba, S. Asur, and B. A. Huberman, "Influence and Passivity in Social Media," in Proceeding of the 22th international conference on World Wide Web - WWW '11, 2011, p. 113-114.