Who Also Likes It? Generating the Most Persuasive Social Explanations in Recommender Systems

Beidou Wang[†♯], Martin Ester[♯], Jiajun Bu[†], Deng Cai[‡]
[†]Zhejiang Provincial Key Laboratory of Service Robot, College of Computer Science, Zhejiang University, China
[♯]School of Computing Science, Simon Fraser University, Canada
[‡]State Key Lab of CAD&CG, College of Computer Science, Zhejiang University, China
[♯]{beidouw,ester}@sfu.ca, [†]bjj@zju.edu.cn,[‡]dengcai@cad.zju.edu.cn

Abstract

Social explanation, the statement with the form of "A and B also like the item", is widely used in almost all the major recommender systems in the web and effectively improves the persuasiveness of the recommendation results by convincing more users to try. This paper presents the first algorithm to generate the most persuasive social explanation by recommending the optimal set of users to be put in the explanation. New challenges like modeling persuasiveness of multiple users, different types of users in social network, sparsity of likes, are discussed in depth and solved in our algorithm. The extensive evaluation demonstrates the advantage of our proposed algorithm compared with traditional methods.

1.Introduction

With over 1.73 billion users around the world, social networking services[1], like Facebook, Google Plus and Twitter, not only capture social relations among people, but also record our preferences towards various items based on our social network activities like post, retweet, like and +1.

This kind of social information is crucial to recommender systems, as previous studies on social influence theories(Cialdini 2001) have already proved that our preferences can be easily impacted by the actions of those around us. For instance, if our friends keep recommending a movie, we are very likely to try it out. Due to this phenomenon, many recommender systems provide extra social information about other people who also like the item, to achieve better performance.

Given an item i to be recommended to user u, we define a social explanation as a statement with the form of "A and B also like the item". Social explanations are widely used in different kinds of recommender systems, including the Facebook page recommendations, Twitter people you may like recommendations and even Google Adwords recommendations, with examples summarized in Figure 1.

The major benefit of social explanation is the *persuasiveness* it brings to the recommendation result. That's to say social explanations convince users to try or buy items(Tintarev and Masthoff 2007). For example, a trusted friend's recommendation may increase our interest to buy an item because we believe in our friend's judgement, or we may want to try something out because we want to talk about it with our friends. Previous studies already confirm that social explanations do help to increase the persuasiveness of the recommendation result (Sharma and Cosley 2013).

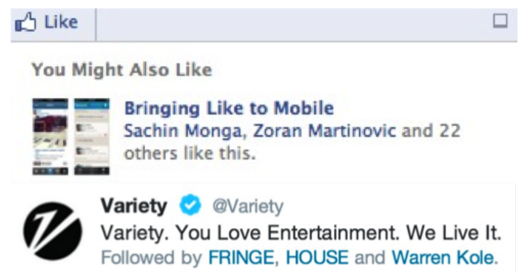

Figure 1: Social Explanations Used in Facebook and Twitter

Despite the wide use of social explanations in systems such as Facebook and Twitter, how to generate the most persuasive social explanation has received little attention in the literature. To the best of our knowledge, this paper proposes the first algorithm to generate the most persuasive explanation , *i.e.* to determine the most persuasive set of users to be put in the explanation.

Our task can be formulated as a ranking problem. Intuitively, we may want to rank users based on user interaction frequency in the social networks and select the top K users for the explanation. However, there are several challenges which make this naive method fail:

Modeling Persuasiveness of multiple users: Is the persuasiveness of multiple users a simple summation of the persuasiveness of individual users? Is adding more users into the explanation always having a positive effect on the persuasiveness? A new model needs to be designed to predict the persuasiveness of multiple users.

Different types of users: In social networks, besides our friends, there are other types of users like celebrities, experts and strangers, which may have different persuasiveness. For instance, when recommending a movie, an explanation like "Your good friends A and B also like it"

[1]http://www.emarketer.com/Article/Social-Networking-Reaches-Nearly-One-Four-Around-World/1009976

will have different persuasiveness than "Your good friend A and president Barack Obama also like it".

Sparsity of likes: Social explanation relies on other people who also like the item. However, in real life we usually have a limited number of social network friends and a huge number of items that need to be recommended, which results in a severe sparsity of like information. In many cases, when we try to generate a social explanation for an item recommendation, it may turn out none of our friends has ever tried it.

Because of the above mentioned challenges, a novel two phase ranking approach is proposed to generate the most persuasive social explanation.

Our research problem can be defined as: for item i to be recommended to user u, determine the most persuasive set of up to k users from the social network S. The social network S is modeled as a directed graph with weighted edges derived from various social network relations and user interactions, and with different types of nodes standing for different types of users (e.g. normal friends, celebrities and experts). Persuasiveness, defined as the power of convincing users to try(Tintarev and Masthoff 2007), is evaluated by real user's rating on the persuasiveness of various explanations, which will be explained in details in section 4 and 5.

Our proposed ranking model is made up of two phases. In the first phase, for recommending an item i to a user u, the persuasiveness score of a single candidate user c is predicted based on a Support Vector Regression (SVR) framework, considering features including social relations, social user interactions, user types and social influence of the candidate. In the second phase, we predict the persuasiveness score of a set of users based on the predicted persuasiveness of the individual users, by taking the marginal net utility of persuasiveness, credibility of the explanation and reading cost into consideration.

The main contributions of the paper are as follows:

1. We are the first to introduce the problem of generating the most persuasive social explanation of a recommendation.
2. We propose a two phase ranking method adopting a machine learning approach to generate the most persuasive explanation, by finding the optimal set of users to be put in a social explanation.
3. The sparsity of like information is addressed by taking into account not only friends, but also celebrities and experts from the social network.

2. Related Work

Explanations in intelligent systems originated in the area of Expert Systems (Andersen et al. 1989)(Buchanan and Shortliffe 1984). Explanations convince a user to accept a recommendation more easily because it provides transparency to the recommender system(Papadimitriou, Symeonidis, and Manolopoulos 2012). Herlocker et al. provided one of the first studies on explanations in recommender systems by evaluating 21 types of interfaces of recommendation explanations and found out that showing the ratings of similar users is the most persuasive explanation(Herlocker, Konstan, and Riedl 2000). Explanations for recommendation are predominantly based on heuristics(Tintarev and Masthoff 2007), but there are also works based on case-based reasoning(Doyle, Tsymbal, and Cunningham 2003) and model based approaches(Lea, Spears, and de Groot 2001).

Explanations can be used for multiple purposes in recommender systems. Tintarev et. al. summarized that there were 7 major aims of recommendation explanations, including transparency, scrutability, trust, effectiveness, persuasiveness, efficiency and satisfaction(Tintarev and Masthoff 2007). Previous studies usually focus on achieving one or a few of above mentioned aims. Trade offs may exist between different aims of recommendation. Bilgic et. al. pointed out that a system optimizing the persuasiveness of explanations may not be helpful for accurately estimating the quality of an item(effectiveness aim)(Bilgic and Mooney 2005). In our paper we only consider the persuasiveness aim of social explanation, because this aim is the most important for increasing user interactions and ad click rates.

With the rapid growth of social network, social networking services now capture not only users' social relations but also users' social interactions, which are both beneficial for recommendation explanation. Many large web service suppliers like Facebook and Netflix, have already started to use social explanations in their recommender systems(Papadimitriou, Symeonidis, and Manolopoulos 2012). Previous studies confirmed that users from our social network friends performed better in social explanations compared with users derived from similar neighbors(Tintarev and Masthoff 2007). Sharma et. al. also confirmed that social explanation did help to increase the persuasiveness of recommendation result based on real user studies on social explanation. However, this work only considers social explanations made up of single users and only test performance of naive strategies like using the close friends in social explanations(Sharma and Cosley 2013). To the best of our knowledge, generating the most persuasive set of users for social explanation has never been studied in previous works.

3. Problem Definition

We formalize the problem of generating the most persuasive social explanation as a ranking problem. Given a set of users U and a set of items I, for recommending an item $i \in I$ to a user $u \in U$, we define a *social explanation* as the statement "Users e_1,e_2,...,e_k also like item i", with given k. We define the candidates C as $C \subseteq U, \forall c \in C, c$ likes i. We define C_k as all the subsets of C with cardinality k, $C_k = \{E \big| E \subseteq C, |E| = k\}$. Since the only difference between two social explanations is the two size-k sets of candidate users, our task is essentially to predict the persuasiveness score for all the elements in C_k and find out the set E with the highest predicted score. Assuming that the predicted score reflects the persuasiveness of the explanation as perceived by the users, we claim our algorithm aims to generate the most persuasive social explanation.

We also define the social network S as a directed graph, in which each node v stands for a user u and each edge e stands for either one type of social relation (*e.g.* follow in

Table 1: Features Used In SVR Model

Category	Used Features
User relation	Whether is a follower, followee, active follower; Common Friends Number
User interaction	Reply, Retweet, Like, Mention (Note these interactions are directed)
User type	User type (experts, celebrities, friends and strangers), Expert/Celebrity Domain(18 large domains)
User Influence	Follower number, average number of reply, like, retweet for each tweet

Twitter) or some social interactions (*e.g.* retweet, mention, reply in Twitter) between two users. The nodes are with different types, corresponding to different types of users (*e.g.* celebrities, experts and friends).

4. Two-Phase Ranking Model for Social Explanation

The key task is to predict the persuasiveness scores for C_k, the size-k sets of candidate users and we use a machine learning approach to train a two-phase ranking model for this task. In the first phase, we predict the persuasiveness score for a single user with a Support Vector Regression (SVR) based model considering various types of features derived from the user's social network. In the second phase, after gaining persuasiveness scores of single users, we adopt them to evaluate the persuasiveness scores of C_k based on our new multiple-user persuasiveness prediction model.

4.1 Phase I: Persuasiveness of a Single User

When recommending item i to user u, an SVR based ranking model is used to predict the persuasiveness score for each candidate user c and the model is made up of two parts: the ranking features for each tuple (u, i, c) and a ranking function to predict persuasiveness $per(u, i, c)$ given the ranking features.

4.1.1 Ranking Features We consider 4 types of features in our SVR model, which is summarized in table 1.

User relation features (UR): We always trust our friend and a friend's endorsement of a recommendation will definitely make it more convincing. We define user relation features $UR(u, c)$ as a boolean value vector with each dimension representing whether one type of relation exists between (u, c).

User interaction features (UI): According to previous studies on social networks, users only communicate with a few of explicitly declared friends(Wang et al. 2013). Fortunately, user interaction features like comments, retweets, likes, enable us to model some implicit networks which can be better indicators of actual social relationships between users. We define user interaction features $UI(u, c)$ as a vector with each dimension denoting one type of user interaction and the value of each dimension is the normalized frequency of the corresponding user interaction between (u, c).

User types features (UT): In our model, users are divided into 4 types, *i.e.* friends, celebrities, experts and strangers. It's especially worth mentioning 2 types of users: celebrities and experts. For celebrities, like movie stars and political leaders, their influence in real life may enable them to cast substantial persuasiveness on a user even if no explicit social relations or interactions exist between them. For experts, like movie critics in movie recommendation, even if they are unknown to the user, their expertise can make their endorsement very persuasive. These two types of users help to significantly expand candidate user set, which helps to solve the sparsity of like data problem.

We define user types feature $UT(c, i)$ as two "1-of-K" vectors. One records the type of candidate c and the other records the domain of c (for celebrities and experts). It's worth noting that whether a user is an expert is related with the domain of the item we recommend.

User Influence features (INF): Intuitively, the more influential a user is, the more trust worthy he looks, which increases his persuasiveness. Moreover, for celebrities, high influence may even offer them the power to bring persuasiveness to users without explicit social relations. We define the social influence feature $INF(c)$ as a set of scalar features, *i.e.* the number of followers, the average number of retweet, reply, like for each tweet.

4.1.2 Ranking Function Many machine learning models can be used as a ranking function to predict the persuasiveness score for a single user. We choose support vector regression (SVR) , because it is a sophisticated proven regression algorithm which is adaptive to complex systems and with a good generalization ability (Wu et al. 2008).

Let's assume that we are recommending an item i to user u considering a candidate user c, we use SVR to compute a score to serve as the persuasiveness score $per(u, i, c)$, which denotes the persuasiveness c brings to the social explanation. We define $x_{u,i,c}$ as the feature vector corresponding to the tuple (u, i, c).

$$x_{u,i,c} = \{UR(u, c), UI(u, c), UT(c, i), INF(c)\} \quad (1)$$

The set of training data is as $\{(x_1, y_1), ..., (x_n, y_n)\}$, where $x_j \subset R^m$ stands for the feature vector for a tuple (u, i, c) in which m is the number of feature dimensions, and $y_j \subset R$ stands for the corresponding persuasiveness value.

A generic SVR estimating function is of the form:

$$f(x) = (w \cdot \phi(x)) + b \quad (2)$$

$w \subset R^m$,$b \subset R$ and ϕ stands for a nonlinear transformation from R^m to high-dimensional space. The core goal of SVR is to learn the value of w and b to minimize risk of regression.

$$Risk(f) = B \sum_{j=0}^{n} L(f(x_j) - y_j) + \frac{1}{2}||w||^2 \quad (3)$$

$L(\cdot)$ is a loss function and B is a constant used to determine penalties to estimation errors which is determined with grids search and cross-validation techniques. We experiment the performance of different kernel functions and choose kernel function with best performance (RBF kernel). Details of SVR can be found in (Smola and Schölkopf 2004).

4.2 Phase II: Persuasiveness of a set of Users

Intuitively, to predict the persuasiveness score for a set of users, we may think about adding up the single-user persuasiveness scores generated in phase I. However, the idea of simple summation of individual users' persuasiveness scores may fail due to the following challenges.

Marginal utility diminishing rules of persuasiveness: We argue that the persuasiveness follows the Law Of Diminishing Marginal Utility. According to the marginal net utility theory(Wang and Zhang 2011), in our task, the *marginal utility* can be interpreted as the additional persuasiveness the social explanation gains after adding an additional candidate into the explanation, which decreases as the number of added users increases. For instance, if we already have several persuasive users, the marginal utility of adding another persuasive user drops. Therefore, it requires us to use a marginal utility model to evaluate the summation of the persuasiveness.

Creditability & reading cost: Does adding more users always lead to a positive effect on the persuasiveness? The answer is negative due to multiple reasons, among which are the decrease of explanation credibility and increase of reading cost, with details explained in section 4.2.2.

Intuitively, we should add users into the explanation in a descending order according to the persuasiveness score we predicts for individual users in phase I. After adding an additional user, even though the summation of persuasiveness increases, the credibility of explanation decreases (because the later added users are always less persuasive) and the reading cost increases. Moreover, the marginal utility of adding additional candidate decreases with the increase of the number of added users. So the best set of candidates requires the best trade off of the summation of persuasiveness, the explanation credibility and the reading cost.

So we propose the following model to predict the persuasiveness score of a set of users. Given recommending item i to user u, for a size-k set of candidates E, the persuasiveness score $P(E, u, i)$ is defined as

$$P(E,u,i) = \alpha log_a \sum_{j=1}^{k} per(u,i,e_j) + \beta \frac{1}{k} \sum_{j=1}^{k} per(u,i,e_j) + \gamma C_k \tag{4}$$

where the first term $log_a \sum_{j=1}^{k} per(u,i,e_j)$ stands for the summation of persuasiveness of a set of individual candidates. The second term denotes the credibility of the explanation. The third term is a constant only related to the total number of candidates in the explanation (k). α, β, γ are variables trained by a standard regression model based on a set of labeled multiple-user social explanation data.

In our framework, we do not need to predict $P(E, u, i)$ for all elements of C_k, which will be time consuming. Details are discussed in section 4.3.

4.2.1 Persuasiveness Summation Based on Cobb-Douglas Utility Function *Marginal utility* is used in economics and marketing research to represent the additional utility when consuming an additional unit of a product or service. *The Law Of Diminishing Marginal Utility* states that the marginal utility of a product or service drops while the consumption of it increases. We argue that the same rule should also be applied to persuasiveness summation, which is to say that the additional persuasiveness the social explanation gains after adding an additional candidate decreases as the number of added user increases.

In our work, we adopt the Cobb-Douglas utility function(Cobb and Douglas 1928). The function is widely used due to its mathematical characteristic: the ability of modeling the diminishing marginal utility. The functional form is:

$$U(E) = log_a \sum_{j=1}^{k} per(u,i,e_j)$$

where a is a variable which can be trained with labeled data.

4.2.2 Credibility of Social Explanation & Reading Cost Due to the sparsity of like information in social networks, there may not always exist enough highly persuasive users for social explanations. Based on our user study, adding less persuasive users (*e.g.* unacquainted candidates) may have a negative effect on the persuasiveness. We believe that this is because it may make the explanation look unreliable as it adds a not so convincing candidate in the social explanation. So we use $\frac{1}{k} \sum_{j=1}^{k} per(u,i,e_j)$, the average of persuasiveness of the k individual users as the credibility level of the social explanation.

Users only have very limited time for reading the social explanation. Adding too many candidates, may distract users from noticing the real persuasive users in the explanation. Moreover, adding too many users in the explanation may make the users feel overwhelmed and decrease their interest in accepting the recommendation. So we add a constant C_k in (4), to represent the reading cost when the social explanation is made up of k different candidates.

4.3 Efficiency

Social networks are made up of millions of users, so strategies are required to guarantee the efficiency of our model.

Single Candidate Persuasiveness Prediction: In phase I of our algorithm, for users who have large numbers of friends in social network, we only consider the top N most interactive friends (N is 10000 in our experiments), because frequently interacting friends are more persuasive compared with infrequent ones. For experts, celebrities and organizations, we only consider the ones with high influence and from the corresponding domain of the recommended item.

Multiple Candidates Persuasiveness Prediction: In phase II of our algorithm, according to our objective function, we do not need to consider every set of candidate users. Instead, we only need to consider the top k candidate users with the highest single user persuasiveness score from phase I. Due to the monotonicity property of our objective function (4) in phase II, $P(E \cup \{e_j\}, u, i) \geq P(E \cup \{e_k\}, u, i)$, if $per(u,i,e_j) \geq per(u,i,e_k)$. So we only need to decide whether we should use the top1, top2,..., top k users to form our explanation.

Sparsity of Relation and Like: For ordinary users, the number of friends in a social network is usually limited, which ensures the efficiency of Phase I. Moreover, due to the sparsity of like data, when recommending item i to u, the number of friends who like i is limited, which ensures the efficiency of our algorithm. The impact of the sparsity will be evaluated in the section 6.

5. Experiments

5.1. Experiment Design & Data Collection

In our experiment, participants are given a recommendation scenario in which we offer recommendations on TV series with various kinds of social explanations. Users rate the persuasiveness of the social explanation in a 7-level Likert scale approach, since many psychometricians advocate using seven levels(Nunnally, Bernstein, and Berge 1967). We then collect the participants' social relation and interaction data from Weibo, the largest micro-blogging system in China. Users related with the participants, celebrities and experts are also extracted from the social network to serve as candidates. To simulate the sparsity of like information, interest in TV series is randomly assigned to all the users we crawled. Participants are trained to only rate on the persuasiveness of social explanation in spite of the actually preference of the recommended item and assume all the social explanations generated are real.

30 participants are invited to generate 3,243 ratings for social explanations (2030 for training and 1213 for testing). We collect 49,183 users from social network as candidates, along with 37,072 user relations and 120,764 interactions.

5.2. Comparison Algorithms

- Random Friends (**RF**). Used as a baseline, in **RF**, when recommending item i to u, we randomly pick k friends of u who also like i to generate the social explanation.
- Good Friends (**GF**). Used as a baseline, in **GF**, when recommending i to u, we rank all candidates who also like i in a descending order of their of interaction frequencies with u and use the top k candidates in the explanation.
- Single User Persuasiveness Contrast Algorithm (**SC**). To test how different types of features affect the precision on predicting the persuasiveness of a single user, we eliminate one type of features at a time from our single user prediction model.
- Multiple User Persuasiveness Contrast Algorithm (**MC**). To test how persuasiveness summation, credibility and reading cost affect the performance of predicting the persuasiveness for a set of users, we eliminate each of them at a time from our prediction model for multiple users.

In all of the baselines and our algorithm, k is set to 3.

Table 2: Algorithm Performance Comparison

Explanation Strategy	Average Rating	Std. Dev	Coverage
Random Friends	4.17	1.81	75.2%
Good Friends	4.46	2.01	97%
2-phase Ranking	**5.07**	1.86	**100%**

Table 5: Influence of Adding More Users

	2 users	2 users+ 1 unacquainted Users
Average Rating	5.7	4.3

6. Results & Analysis

6.1 Algorithm Performance Comparison

An A/B test is conducted to compare our algorithms with baselines and the result is shown in table 2. Based on user's average rating on the persuasiveness of the social explanation, our algorithm shows a 21% improvement compared with RF and a 14% improvement compared with GF. The result is in accordance with our expectation. GF outperforms RF, as besides user's social relations, GF also takes the frequency of interactions into consideration and intuitively close friends will be more persuasive than ordinary friends. While our algorithm performs the best, which benefits from the exploitation of all kinds of additional features from social network, like user types and user influence and a careful designed persuasiveness model for a set of users.

It's also worth noting that both RF and GF cannot generate social explanations for all the recommendation due to the sparsity of like data, for instance, in scenarios that none of a user's friends has ever tried an item. However, our 2-phase ranking algorithm can still provide social explanations in these situation, by adding persuasive celebrities and experts into explanations. We define the percentage of recommendations in which one strategy can be used to generate the social explanation as *coverage*, with results in Table 2.

6.2 Feature Importance Evaluation

Based on table 3, all of the 4 types of features we considered do help to improve the precision of persuasiveness score prediction for a single user. Social relations and social interactions are the most important 2 types of features. It's a little surprising that social relation features are more important than the social interaction features and we propose two possible reasons. First, since we extract our data from a micro-blogging system, a user's interactions like retweet and reply with a candidate c may only mean that he is interested in certain posts published by c, instead of acquainted with c or trusting c. Second, spam accounts like to reach out to users actively in micro-blogging systems, which may result in noise in the user interaction data.

Based on table 4, all the three factors we considered in modeling the persuasiveness of a set of users contribute to the improvement of accuracy in prediction and the credibility factor is of the most significance.

Table 3: Feature Comparison for Single User Persuasiveness Prediction

Evaluation Measure	SC_withAll	SC_withNoRelation	SC_withNoInteraction	SC_withNoTypes	SC_withNoInfluence
MSE	2.34	2.99	2.47	2.40	2.37

Table 4: Feature Comparison for Predicting Persuasiveness for a set of Users

Evaluation Measure	MC_withAll	MC_withNoPersuasivnessAggregation	MC_withNoCredibility	MC_withNoReadingCost
MSE	2.03	2.07	2.16	2.05

Table 6: Average Candidate Set Size Comparison

	Random Friends	Good Friends	2-phase Ranking
Size	24	130	330

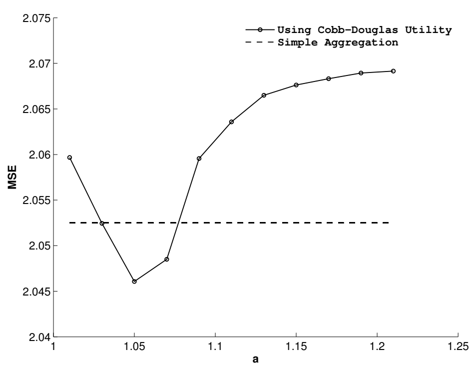

Figure 2: Marginal Utility Diminishing Test

6.3 Marginal Utility Diminishing Test

In section 4, we argue that the summation of the persuasiveness for multiple candidates should follow Marginal Utility Diminishing Law, so a contrast experiment is designed to test it. For the persuasiveness summation term in equation (4), we compared using Cobb-Douglas Utility Function with simply summing the persuasiveness score of all the individual candidates. As shown in Figure 2, the best performance is achieved by adopting Cobb-Douglas Utility Function with parameter a set to 1.05, which confirms our argument.

6.4 Influence of Adding More Users

To test whether adding more users will always do no harm to the social explanation, a user study is designed. We first generate a set of explanations made up of two persuasive users and we then add an unacquainted user into each explanation. We ask users to rate these explanations and as shown in table 5, a 25% decrease of the persuasiveness rate occurs after adding the unacquainted user, which further confirms our assumptions in section 4.2.2.

6.5 Impact of Sparsity

For ordinary users, the average number of friends is relatively small compared with total number of users from social network. In our data set, the average number of friends is 479 and the total number of experts and celebrities we considered is 1792, which ensures the efficiency of computing the score for all individual candidates in Phase I.

The sparsity of like information will result in problems of the coverage of the recommendation. Traditional social explanation strategies (e.g. *GF*, *RF*) that consider only explicit friends often cannot generate a valid explanation because none of the user's friends has ever tried the item. In our approach, by incorporating the celebrities and experts without explicit user relations, the candidate set is effectively expanded. According to table 6, our algorithm almost triples the average size of the candidate set compared with *GF* .

However, it is worth noting that even though our algorithm has greatly expanded the candidate set, due to the sparsity of likes, the average size of candidate set (330) is relative small, which further improves the efficiency of Phase II.

7. Conclusion

In this paper, we present the first algorithm to generate the most persuasive social explanation by recommending the optimal set of users to be put in the explanation. A two-phase ranking algorithm is proposed and inside our algorithm, we present the first model for predicting the persuasiveness of a set users, taking factors like marginal utility of persuasiveness, credibility of explanation and reading cost into considerations.

We find our algorithm gains the best performance compared with all the baselines and the sparsity of like is solved by considering different types of users, like celebrities and experts, who can bring persuasiveness to a user even if without explicit social relations between them. Based on our experiments, we also confirm that the aggregation of persuasiveness follows the marginal utility diminishing laws and more users do not necessarily lead to more persuasiveness.

Many future works can be further explored. For instance, what if the predicted score of recommendation result is not in accordance with the persuasiveness of the social explanation? How to cope with the trade off of the two factors to achieve the best result can be an interesting future research topic.

Acknowledgment

This paper is supported by the National Basic Research Program of China (973 Program) under Grant 2013CB336500, the National Science Foundation of China (61173186) and the National Key Technology R&D Program(2012BAI34B01).

References

Andersen, S. K.; Olesen, K. G.; Jensen, F. V.; and Jensen, F. 1989. Hugin-a shell for building bayesian belief universes for expert systems. In *IJCAI*, volume 89, 1080–1085.

Bilgic, M., and Mooney, R. J. 2005. Explaining recommendations: Satisfaction vs. promotion. In *Beyond Personalization Workshop, IUI*, volume 5.

Buchanan, B. G., and Shortliffe, E. H. 1984. *Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project (The Addison-Wesley series in artificial intelligence)*. Addison-Wesley Longman Publishing Co., Inc.

Cialdini, R. B. 2001. Influence: Science and practice. *Boston: Allyn & Bacon*.

Cobb, C. W., and Douglas, P. H. 1928. A theory of production. *The American Economic Review* 18(1):139–165.

Doyle, D.; Tsymbal, A.; and Cunningham, P. 2003. A review of explanation and explanation in case-based reasoning. *Dublin, Trinity college https://www. cs. tcd. ie/publications/tech-reports/reports* 3.

Herlocker, J. L.; Konstan, J. A.; and Riedl, J. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 241–250. ACM.

Lea, M.; Spears, R.; and de Groot, D. 2001. Knowing me, knowing you: Anonymity effects on social identity processes within groups. *Personality and Social Psychology Bulletin* 27(5):526–537.

Nunnally, J. C.; Bernstein, I. H.; and Berge, J. M. t. 1967. *Psychometric theory*, volume 226. McGraw-Hill New York.

Papadimitriou, A.; Symeonidis, P.; and Manolopoulos, Y. 2012. A generalized taxonomy of explanations styles for traditional and social recommender systems. *Data Mining and Knowledge Discovery* 24(3):555–583.

Sharma, A., and Cosley, D. 2013. Do social explanations work?: studying and modeling the effects of social explanations in recommender systems. In *Proceedings of the 22nd international conference on World Wide Web*, 1133–1144. International World Wide Web Conferences Steering Committee.

Smola, A., and Schölkopf, B. 2004. A tutorial on support vector regression. *Statistics and computing* 14(3):199–222.

Tintarev, N., and Masthoff, J. 2007. A survey of explanations in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, 801–810. IEEE.

Wang, J., and Zhang, Y. 2011. Utilizing marginal net utility for recommendation in e-commerce. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, 1003–1012. ACM.

Wang, B.; Wang, C.; Bu, J.; Chen, C.; Zhang, W. V.; Cai, D.; and He, X. 2013. Whom to mention: expand the diffusion of tweets by@ recommendation on micro-blogging systems. In *Proceedings of the 22nd international conference on World Wide Web*, 1331–1340. International World Wide Web Conferences Steering Committee.

Wu, X.; Kumar, V.; Ross Quinlan, J.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.; Ng, A.; Liu, B.; Yu, P.; et al. 2008. Top 10 algorithms in data mining. *Knowledge and Information Systems* 14:1–37.