

Exploiting Game Theoretic Analysis for Link Recommendation in Social Networks

Tong Zhao^{1,2}, H. Vicky Zhao³, Irwin King^{1,2}

¹Shenzhen Key Laboratory of Rich Media Big Data Analytics and Applications, Shenzhen Research Institute,

The Chinese University of Hong Kong, Shenzhen, China

²Department of Computer Science and Engineering,

The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

³Electrical and Computer Engineering Department, University of Alberta, Edmonton, AB Canada
{tzhao, king}@cse.cuhk.edu.hk, vzhao@ece.ualberta.ca

ABSTRACT

The popularity of Online Social Networks (OSNs) has attracted great research interests in different fields. In Economics, researchers use game theory to analyze the mechanism of network formation, which is called Network Formation Game. While in Computer Science, much effort has been done in building machine learning models to predict future or missing links. However, there are few works considering how to combine game theoretic analysis and machine learning models. Therefore, in this paper, we study the problem of Exploiting Game Theoretic Analysis for Link Recommendation in Social Networks. Our goal is to improve link recommendation accuracy via leveraging the power of Network Formation Games into machine learning models. We present two different approaches to solve this problem. First, we propose a three-phase method that straightforwardly combines game theoretic analysis with machine learning models. Second, we develop a unified model, *BPRLGT*, that incorporates Network Formation Game into a Bayesian ranking framework for link recommendation. Specifically, *BPRLGT* takes advantage of network topology and we design a game theoretic sampling approach to improve its training process. The experiments are conducted on four real world datasets and the results on all datasets demonstrate that both our proposed three-phase method and the unified ranking model outperform the baseline methods.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

General Terms

Algorithms, Experimentation, Performance

Keywords

Game Theory; Network Formation Game; Link Prediction

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CIKM'15, October 19–23, 2015, Melbourne, Australia.

© 2015 ACM. ISBN 978-1-4503-3794-6/15/10 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/2806416.2806510>.

1. INTRODUCTION

Online social networks (OSNs) play an important role for people to satisfy their social needs and seek novel information on Internet. As shown by recent studies, there are over 5.7 billion users on the world's top 21 social networks and 72% of all Internet users are active on social media for sharing feelings, posting pictures and most importantly, making friends. Such popularity has attracted great research interests from different fields.

In Economics, a popular research direction is to use game theory to study the mechanism of the formation of social networks, which refers to *Network Formation Game* (NFG). Typically, in NFG models, utility functions are defined to represent how much utility people can gain from their connected neighbors and the models assume that people build social links in order to maximize their utility. Here utility represents the benefit or profit for people. Based on NFG, economic researchers try to investigate the following problems: 1) given a specific social network, how to find the particular utility functions that can explain why people form such a network; 2) how different definitions of utility function can trigger people to form different social networks in order to maximize their utility. To answer these questions, several possible modeling approaches are proposed, including simultaneous move modeling, pairwise-stability modeling [16], bargaining for network formation [8], dynamic modeling [31], farsighted network formation modeling [11], etc.. However, most of these works focus on explaining the formation of current network and cannot be directly used for link prediction.

While in Computer Science, much effort has also been done for social network analysis, especially in predicting new links for people. Many methods are proposed to solve this problem. Several atomic propagation models are proposed in [12] such as direct propagation, co-citation propagation, transpose propagation and coupling propagation for trust link prediction. Leskovec et al. [21] leverage the topological feature of a social network to predict the trust and distrust relations among users. Tang et al. propose a method for link prediction by capturing the evolution of trust in [34]. Guha et al. propose a framework for modeling trust propagation [12]. Dong et al. [10] design a factor graph method to predict trust links between users across heterogeneous networks. While the existing methods can learn the network and predict links to users, none of them consider how users' utility change when building the recommended links.

Therefore, in this paper, we study the problem of *Exploiting Game Theoretic Analysis for Link Recommendation in Social Networks*. The goal is to improve link recommendation (link pre-

diction) accuracy via leveraging the power of Network Formation Games (NFG). However, there are several challenges when combining game theoretic models and link prediction models. First, most of the NFG models focus on explaining the formation of current networks and assume that no users can further improve their utility through building new links. This assumption conflicts the goal of link prediction models. Second, the goals of NFG analysis and link prediction models are different. NFG models aim to maximize users' utility value while the goal of link prediction models is to maximize the accuracy of link recommendation. To our best knowledge, there is no work investigating the connections between the two goals. Third, no existing link prediction models consider the utility analysis when training, how to integrate utility analysis into model training is a challenging problem.

To address these challenges, in this paper, we first conduct an empirical study to investigate how well NFG models describe the formation mechanism of real networks. Then we propose a three-phase approach that leverages the power of game theory into link prediction models. Given a social network G , the first phase employs NGF models to calculate an original utility value of each user in G ; at the second phase, link prediction models are used to find the candidate links; after adding the candidate links to G , the third phase then calculates users' utility via NFG models again and uses the utility difference based on the current and original utility as a criterion to help re-rank the candidate links. In particular, the links which are highly-ranked by link prediction models but cannot improve the utility of connected users will be re-ranked to a lower priority for final recommendation.

Although the three-phase method provides a simple way to combine NFG models and machine learning models, it processes the knowledge from NFG models and machine learning models in separated steps. In particular, NFG models have no contribution when training machine learning models and no utility information is considered when generating candidate links.

To alleviate these problems and better integrate NFG models into machine learning models, we further propose a novel unified pairwise ranking model, namely, *BPRLGT* (*Bayesian Personalized Ranking for Link recommendation with Game Theory*), integrating NFG models into a Bayesian ranking model for link recommendation. Our model is built based on Bayesian Personalized Ranking framework (BPR) [27], which always requires sampling technique to select training data. In BPRLGT, we emphasize on designing a game theoretic sampling approach to improve model training. More specifically, different from uniformly sampling and other heuristic sampling approaches [29] such as popularity-based sampling, etc., the proposed game theoretic sampling approach considers users' utility and helps the model punish the links that cannot increase users' utility but with high-rank. Moreover, BPRLGT model also takes network topology into account, which makes BPRLGT more suitable for link prediction.

We conduct experiments on four real world social network datasets and the results show that our proposed three-phase method and the unified model consistently outperform other state-of-the-art methods.

Organization. In Section 2, we discuss the related work. In Section 3, we provide the background and then formulate the problem; Section 4 details the proposed three-phase method and the unified ranking model with the game theoretic sampling technique. Section 5 demonstrates the experimental results and verifies the contribution of integrating game theoretical analysis into link prediction problems. Finally, Section 6 concludes the paper.

2. RELATED WORK

In this section we review the related work, regarding the Network Formation Game, link prediction in social network, as well as one-class recommendation problem.

Network Formation Game. Recent years, Game Theory [5, 17, 18, 22] has been used to investigate people's behaviors and social relations in social networks. Typically, a utility function is defined to represent people's utility in a social network. By analyzing parameters in utility function and the network structure, researchers try to explain the mechanism of the network formation. We refer readers to [14] and [15] for a detailed survey of models in Network Formation Games. Xiao et al. [35] formulate the wireless security problem as an indirect reciprocity game, and propose a method that applies the indirect reciprocity principle to suppress attacks in wireless networks. In [16], the authors study the stability and efficiency of social and economic networks. Bramoulle et al. [7] provide a general network analysis to a wide class of games, including strategic innovation, public goods and social interactions. Mele [23] analyzes the network formation by a dynamic meeting protocol and Badev [2] further extends the model. Paula et al. [26] provide a framework to learn parameters of utility function in network formation models.

Link Prediction. A large number of studies about link prediction [9, 36, 32, 1, 34] have been proposed based on either similarity analysis or propagation analysis. Siegler et al. point out that there is a strong and significant correlation between trust link and similarity [39] and then use the similarity for link prediction. Guha et al. propose a framework for modeling trust propagation [12]. Leskovec et al. utilize the topological feature of a social network to predict the trust and distrust relations among users [21]. Tang [32] et al. exploit homophily effect for link prediction. Lee et al. [20] develop an integer programming problem for the link prediction problem. Besides the machine learning models, some game theory-based methods are also proposed for link prediction, but they all need some additional information, such as users' rating or tagging behavior on items, to help define the game parameters. Zappella et al. [37] present a Graph Transduction Game for link prediction and Beigi et al. [6] also employ a game theoretic approach for community detection and link prediction. Few work considers to integrate game theory models into machine learning models to improve the accuracy on link prediction.

One-class Recommendation. One-class recommendation focus on solving recommendation problems when one can only observe some forms of users' implicit feedback from data, typically binary feedback, such as the webpages they browse, the music they listen to, the link they click or whom they make friend with. We view users' observed feedback as positive feedback and others as negative feedback. In this paper, we cast link prediction as an one-class recommendation problem since we can only observe users' positive feedback (existing links) and aim at predicting missing or future links. Several works have studied one-class recommendation and can be mainly divided into two branches: pointwise methods [13, 25] and pairwise methods [27, 30]. Pointwise methods aim to fit a numeric value associated with each evaluated item, while pairwise methods focus on modeling the order, or ranking of the feedback. Pairwise methods always consider implicit feedback as relative relationships indicating that users show higher preference on positive feedback than on negative feedback. In [27], Rendle et al. propose a bayesian personalized ranking (BPR) framework. Following this, various ideas have been proposed that incorporate different types of contextual information into the BPR framework. [19] extends the BPR framework to model both users' feedback on items and on their social relations. In [38], Zhao et al. leverage social connec-

tions to improve item recommendations by building a new preference system. Since the learning method used in most related works is stochastic gradient descent (SGD), how to select the training instances has great impact on model performance. In [29], the authors propose two strategies to sample negative feedback and also design an efficient implementation to improve the convergence of the learning algorithm.

Summary. Our proposed methods differentiate themselves from all these existing methods in that the study of incorporating game theoretic analysis into machine learning models is unexplored in previous work.

3. BACKGROUND AND PROBLEM DEFINITION

In this section, we first present several notations and definitions used in the paper and then detail two Network Formation models, Co-author model and Linear Best Reply model, which will be used in our proposed methods. We analyze the property of Network Formation models and then formally define the problem of *Exploiting Game Theoretic Analysis to Improve Link Recommendation in Social Networks*.

3.1 Definitions and Notations

Let $G = (V, E)$ denote the social network, where V is a set of $|V| = M$ users and $E \subset V \times V$ is a set of links between users. G can be either directed or undirected. N_i represents the directed neighbors of user i in graph G . $e_{ij} \in E(G)$ means there is a direct link between user i and j in G and $e_{ij} \notin E(G)$ represents that there is no direct link between the two users. Given a network G and $e_{ij} \notin E$, let $G + e_{ij}$ denote the network formed by adding the link e_{ij} to the network G , and G/e_{ij} denote the network formed by deleting the link e_{ij} from the network G . \mathbf{I} denotes the identity matrix, $\mathbf{1}$ denotes a vector of ones, A^T denotes the transpose of matrix A and $\det(A)$ represents the determinant of matrix A . We also have the following concept definitions.

Definition 1. Action: Actions in this paper are defined as user's behaviors to build social relations with others. In particular, given G , each user i 's action can be represented as $x_i \in \mathbb{R}^M$, where the j th entry in x_i equals to 1 if $e_{ij} \in E$ and 0 otherwise. Let $\mathbf{x} = \langle x_1, x_2, \dots, x_M \rangle$ denote the action vector for all users and $\mathbf{x}_{-i} = \langle x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_M \rangle$ denote the action vector of all users other than i .

Definition 2. Individual utility: The individual utility represents the profit, quantity or "utility" in result of user's actions. Typically, it can be constructed as the value of a utility function, $U(\cdot)$. In this paper, we assume that users' utility will change only when G changes and only building new links or deleting existing links will affect users' utility. Let $U(x_i; \mathbf{x}_{-i}, G)$ denote the utility function to calculate user i 's utility value in network G since G is formed by action vector \mathbf{x} . For simplicity, we use $U_i(G)$ or U_i to represent user i 's utility value in social network G .

Definition 3. Utility difference: Utility difference of user i , Δ_i , describes the difference when user i builds some new links. Specifically, given G and a set of recommended links, $\{e_{ij}\} \notin E(G)$, $\Delta_i = U_i(G + \{e_{ij}\}) - U_i(G)$. If $\Delta_i \geq 0$, we conclude that e_{ij} is a *beneficial link* for user i . In particular, for a recommended link $e_{ij} \notin E(G)$, if $\Delta_i \geq 0$ and $\Delta_j \geq 0$, we call e_{ij} a *mutual consent link*.

Table 1: Statistics of the Datasets after filtering.

	Ciao	Epinions	COL	FB
#Users	2342	9049	40421	63732
#Edges	51789	320150	316246	1071968
#Avg.Edges	22	35	8	17
Density	9.4×10^{-3}	3.9×10^{-3}	1.9×10^{-4}	2.6×10^{-4}

In game theory, the most important concept is Nash Equilibrium, which describes a solution concept of a Network Formation Game such that no player in the game can gain more by changing only their own actions. Here we also introduce the Nash Equilibrium concept used in Network Formation Games.

Definition 4. Nash Equilibrium: An action vector \mathbf{x} is a Nash Equilibrium if and only if for all users i ,

$$U_i(x_i; \mathbf{x}_{-i}, G) \geq U_i(x'_i; \mathbf{x}_{-i}, G).$$

This definition shows that for user i , his/her utility cannot increase by changing his/her action from x_i to x'_i . In fact, Nash Equilibrium describes the optimal utility that users can gain in network G . Based on the definition of action, we can see that the network G is the result of users' actions and given a network G , we can calculate users' utility based on $U(\cdot)$. In this paper, we aim to improve link recommendation accuracy by considering the utility that users will gain from building the recommended links. Specifically, given a network G and a set of recommended links $\{e_{ij}\}$, we first assume that both G and $G + \{e_{ij}\}$ reach Nash Equilibrium, then we can obtain the optimal utilities, $U_i(G)$ and $U_i(G + \{e_{ij}\})$, for each user i on G and $G + \{e_{ij}\}$. More importantly, we can then judge whether $\{e_{ij}\}$ should be recommended to people based on $\Delta_i = U_i(G + \{e_{ij}\}) - U_i(G)$: if $\Delta_i \geq 0$ for user i , then we conclude that $\{e_{ij}\}$ should be recommended; otherwise, we should not recommend them to people since people cannot increase their utilities by building them.

3.2 Network Formation Games

In this section, we first formally define the Network Formation Game and then present several utility functions that are widely used in NFG. Following that, we analyze their properties when reaching Nash Equilibrium.

A Network Formation Game (NFG) in game theory is always used to analyze user's discretion when they build links in social network and it can be represented by a triple $(G, \mathbf{U}, U(\cdot))$, where $\mathbf{U} = \langle U_1(G), \dots, U_i(G), \dots, U_M(G) \rangle \in \mathbb{R}^M$ records users' utilities in G and $U(\cdot)$ is the utility function which defines how utilities are allocated across the users when they build a particular network. The key component of NFG is to find a proper utility function and ensure that it could explain why users build such a particular network. Specifically, the utility function describes the utilities that users directly receive, taking both benefit and cost of maintaining their social links into account. Different utility functions lead to different models in NFG and also motivate users to set up different network structures. The goal of NFG is to analyze users' utilities when assuming the given network G reach a Nash Equilibrium. In the following subsection, we will introduce two popular NFG models, Co-author model and Linear Best-Response model in details.

3.2.1 Models in Network Formation Games

The Co-author Model. The Co-author model [16] is firstly proposed to describe the collaboration relationship between researchers. Each user in this model is assumed to own a fixed

Input: A network $G = (V, E)$ and δ
Output: Effort vector $\mathbf{U}(G)$
 $\mathbf{U}_S(G) = \mathbf{0}$;
repeat
 Select a subset of users, $S \subseteq V$;
 Construct $G_S(G)$ and G_{V-S} ;
 if $\det(\mathbf{I} + \delta G_S)$ equals to 0 **then**
 Continue;
 end
 $\mathbf{U}_S(G) = (\mathbf{I} + \delta G_S)^{-1} \mathbf{1}$;
until $\mathbf{U}_S(G) \geq \mathbf{0}$ and $\delta G_{V-S} \mathbf{U}_S(G) \geq \mathbf{1}$;
Return $\mathbf{U}_S(G)$.

Algorithm 1: Utility calculation in LBR. G_S denotes the subgraph of G connecting the selected users S and G_{V-S} denotes the graph of links connecting users S to the rest users. $\mathbf{U}_S(G)$ denotes the utility vector when users S are selected to estimate the Nash Equilibrium in network G .

amount of time to spend or pay, so the time user i spends on each collaboration is inversely related to the number of links he/she has, say $n_i = |N_i|$. The utility of user i is represented by

$$U_i(G) = \sum_{j: e_{ij} \in E(G)} \left(\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j} \right). \quad (1)$$

Here we can see that the more links user i has, the lower synergy he/she could gain from this link. Note that there is no direct cost modeling, but the diluted gains reflect the cost when user builds too many links with others. Given G , users' utilities can be directly calculated from Eq (1). In this way, when considering to build new links, user i needs to think about the utility difference after they constructing these links. Specifically, when user i considers to make friends with user k , the utility difference can be estimated as

$$\begin{aligned} \Delta_i &= \frac{1}{n_k} + \frac{1}{n_k(n_i + 1)} + \left(1 + \frac{1}{n_i + 1}\right) \sum_{j: e_{ij} \in E(G)} \frac{1}{n_j} \\ &\quad - \left(1 + \frac{1}{n_i}\right) \sum_{j: e_{ij} \in E(G)} \frac{1}{n_j}, \\ &= \frac{1}{n_k} + \frac{1}{n_k(n_i + 1)} - \frac{1}{n_i(n_i + 1)} \sum_{j: e_{ij} \in E(G)} \frac{1}{n_j}. \end{aligned} \quad (2)$$

If $\Delta_i \geq 0$, we conclude e_{ik} as a beneficial link to user i since it could increase i 's utility.

The Linear Best-Response Model. In Linear Best Response (LBR) model [7], vector \mathbf{U} is used to represent users' effort on network G , \bar{u} is the required efforts that each user needs take in isolation; for each user i , his/her effort can be calculated as following,

$$U_i(G) = \begin{cases} \bar{u} - \delta \sum_{e_{ij} \in E(G)} U_j(G) & \text{if } \bar{u} > \delta \sum_{e_{ij} \in E(G)} U_j(G) \\ 0 & \text{if } \bar{u} \leq \delta \sum_{e_{ij} \in E(G)} U_j(G), \end{cases} \quad (3)$$

where δ is a measure of how much user i 's directed neighbors affect his/her efforts. Thus, user i achieves his/her required efforts \bar{u} through combining his/her own effort and his/her friends' efforts. From this definition, we find that the more efforts user i 's neighbors could offer, the less effort user i herself would take. In LBR model, each user wants to minimize his/her own efforts while reaching the required \bar{u} by others.

LBR model captures the property of several exemplary games including network formation game. For example, researchers benefit from the results of experiments done by their collaborators, which

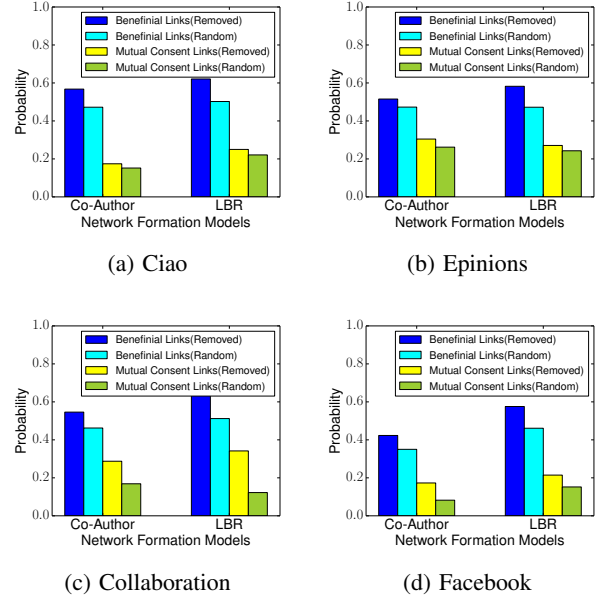


Figure 1: Empirical analysis of Network Formation Models. We investigate how much different network formation models can capture the real network evolution from benefit perspective. We include a random setting for clear comparison. Y-axis represents the probability that Network Formation Game models precisely capture the beneficial links.

explains the formation of academic collaboration networks; people in Facebook learn more by the sharings of their friends, which explains the formation of social networks.

Accordingly, user i 's utility function then can be equivalently represented as

$$U(x_i; \mathbf{x}_{-i}, \delta, G) = -\frac{1}{2} U_i^2 - \delta \sum_{e_{ij} \in E(G)} U_i U_j + \bar{u} U_i. \quad (4)$$

Equation (3) can be obtained by calculating the derivation of Eq (4). However, one can see that from Eq (4), utility value cannot be directly calculated since Eq (4) contains both U_i and U_j . Therefore, we need to borrow the concept of Nash Equilibrium to obtain the final utility of each user. Several works are done for the analysis of Nash Equilibrium of LBR model [3, 4]. According Proposition 1 in [7], we can obtain the utility of each user via Algorithm 1.

Similar to Co-author model, here we also use the utility difference as criteria to identify beneficial and mutual consent links. Specifically, when user i considers to make friends with user k , the utility difference can be estimated as $\Delta_i = -(U_i(G + e_{ik}) - U_i(G))$. If $\Delta_i \geq 0$, we say e_{ik} a beneficial link to user i since it could decrease i 's effort.

3.2.2 Empirical Analysis on Real Networks

As we discussed in Section 1, though the above game theoretic models show their strength on analyzing the formation mechanism of a static network from theory perspective, an important issue still remains: could these model capture the mechanism of evolution of real networks? In particular, we are interested in understanding: **Q1**, if we exactly know the links users will build and apply NFG models to estimate their utilities before and after they setting up these links, whether their utilities would increase accordingly? **Q2**, how precise the models can achieve in identifying beneficial or

mutual consent links? **Q3**, could users gain more utility by adding arbitrary links?

To answer questions **Q1** to **Q3**, we conduct an empirical analysis on four real-world datasets, Ciao, Epinions¹, Collaboration Network (COL)² and Facebook³. The statistics of these datasets are shown in Table 1 and we will introduce the datasets in detail in Section 5. To answer these questions, we first remove $X\%$ links $\{e_{ij} \in E(G)\}$ as \mathcal{N} from network G and note the network with the rest links as $G_{removed}$, then we utilize the models of NFG to calculate the utilities of each user in $G_{removed}$, say $U(G_{removed})$; following that, we add the removed links back and again calculate the utilities of each user, say $U(G)$; thus, for each user i who has link changes, we calculate Δ_i . Based on the description in *Definition 3*, we can then count the number of beneficial links and further obtain the number of mutual consent links. The more a NFG model can identify beneficial and mutual consent links, the better this model will be when using it to model real network evolution. In addition, we also calculate users' utility difference by randomly adding links to $G_{removed}$ for answering **Q3**. The number of randomly added links is same as removed links. We test X as 10, 20 and 30. Since the performance are similar, here we just report the result of $X=20$.

By cross validation, we find the best performance of each models as Figure 1. Here we set $\delta = 0.25$ and $\bar{u} = 1$ in LBR model. We summarize the results by following observations:

- On all datasets, NFG models achieve almost 60% precision on average for capturing beneficial links;
- Co-author model estimates most of the removed links as beneficial links. LBR model always achieves better performance on capturing mutual consent link.
- Compared with the random setting, NFG models have higher probability to model *removed links* as beneficial and mutual consent links than randomly added links. Since the *removed links* are the real links users would like to build, this result verifies NFG models' effectiveness on explaining users' actions in social networks: users are more likely to build new links such that they can increase their utility from them.

In summary, these results indicate that NFG models can partially describe the evolution of real networks from utility perspective, which motivates us to consider an approach to leverage NFG models to help improve the recommendation performance of link prediction methods. In the following section, we will first introduce a three-phase method. Furthermore, we present a unified model that incorporates Network Formation Game into a Bayesian ranking model for link recommendation.

3.3 Problem Formulation

Given the definitions and empirical analysis above, now we can formally define the problem solved in this paper.

Problem 1. Exploiting Game Theoretic Analysis for Link Recommendation in Social Networks: Given a network G , a utility function $U(\cdot)$, the goal is to learn a predictive model from $e_{ij} \in E(G)$ and $U(\cdot)$ such that the model can recommend new links to each user as accurate as possible.

¹Ciao and Epinions datasets can be found from [33]'s first author webpage

²<http://www-personal.umich.edu/mejn/netdata/>

³<http://konect.uni-koblenz.de/networks/facebook-wosn-links>

Input: Recommended link number T , a social network $G = (V, E)$, link prediction model \mathcal{L} and a utility function $U(\cdot)$
Output: A set of recommended links $e_{ij} \notin G$
 Calculate the utility of each user i via $U_i(G)$.
 Learning the link prediction model, \mathcal{L} , on G .
 Rank the probability of the links $e_{ij} \notin G$ based on \mathcal{L} .
 Adding links with top- $2T$ highest probability, $\{e_{ij}\}$, to G .
 Calculate the utility of each user i via $U_i(G + \{e_{ij}\})$.
for each link $e_{ij} \notin E(G)$ **from the top on the ranking list do**
 if $\Delta_i \leq 0$ **and** $\Delta_j \leq 0$ **then**
 Replace the probability of link e_{ij} with the closest following link $e_{i'j'}$, satisfying $\Delta_{i'} \geq 0$ and $\Delta_{j'} \geq 0$.
 end
end
 Re-rank the probability of links.
 Return Top- T links.

Algorithm 2: A game theoretic framework for link prediction.

4. PROPOSED METHODS

In this section, we first introduce a three-phase method that employs NFG models to improve link recommendation performance. The proposed framework is quite general and the idea here is straightforward that by using NFG models to calculate the utility difference when adding links to the network, the links which are highly-recommended by link prediction models but cannot improve the utility of connected users will be re-ranked to a lower priority for final recommendation. Besides this three-phase approach, we also propose a novel unified pairwise ranking model, namely, *BPRLGT*, integrating NFG models into a Bayesian ranking model for link recommendation. Our model is built based on Bayesian Personalized Ranking framework (BPR) [27], which always requires sampling technique to select training data. In *BPRLGT*, we extend the preference assumption of BPR by considering network topology and further propose a game theoretic sampling approach to improve model training.

4.1 A Three-phase Game Theoretic Method for Link Prediction

In this section, we introduce the proposed method that leverages the power of game theory to improve the accuracy of link recommendation in social networks.

Given a network graph G , a utility function $U(\cdot)$, a link prediction model \mathcal{L} , our goal is to recommend T links for users. The proposed method includes three phases. At first phase, we utilize $U(\cdot)$ to calculate the current utility of each user i , as $U_i(G)$ on network G . Then a link prediction method \mathcal{L} is trained on G and suggests a list of new links $\{e_{ij}\}$ for users. After getting the recommended links from \mathcal{L} , at the beginning of third phase, we add them to G and form a new graph, $G + \{e_{ij}\}$. Based on this new graph, we can compute a new utility of each user by calling $U(\cdot)$ again, say $U_i(G + \{e_{ij}\})$. By comparing $\Delta_i = U_i(G + \{e_{ij}\}) - U_i(G)$ with 0, we can identify the beneficial links and further adjust the ranking of suggested links. In particular, we can decrease the ranking of those non-beneficial links. Thus, after the third phase, the links in the top ranking positions are approved by both machine learning method and NFG model. In other words, these recommended links not only have high probability to be set up but also are beneficial links once they could be set up. The algorithm for this framework is shown as Algorithm 2.

Although the proposed framework provides a straightforward way to combine NFG models and machine learning models, and it can improve the recommendation performance as we will show in Section 5, we notice that the framework processes the knowl-

edge from NFG models and machine learning models in separated steps, which makes NFG models insufficient in improving machine learning models directly. To better integrate NFG models into machine learning models, we therefore develop a unified Bayesian ranking model with game theoretic sampling techniques for link recommendation.

4.2 A Unified Game Theoretic Ranking Model for Link Recommendation

Since the observed network data only contains binary feedback between two users (1 for an observed link, 0 for no link), the link recommendation problem then can be cast as a one-class recommendation problem, as we introduced in Section 2. Thus, we propose a pairwise ranking method, *BPRLGT* (*B*ayesian *P*ersonalized *R*anking for *L*ink recommendation with *G*ame *T*heory), to solve this problem. Our proposed model is built upon Bayesian Personalized Ranking framework, BPR [27], which requires sampling technique to select pairwise data for Stochastic Gradient Descent (SGD) training. In BPRLGT, we emphasize on designing a game theoretic sampling approach to improve model training. Moreover, BPRLGT model also takes network topology into account, which makes BPRLGT more suitable for link prediction than BPR.

Model Assumption. We first introduce the basic assumption used in BPR [27] in a link recommendation scenario as following,

$$p_{ui} \succeq p_{uj}, \text{ for } e_{ui} \in E(G) \text{ and } e_{uj} \notin E(G), \quad (5)$$

where p_{ui} represents a user u 's preference on building a link with user i and the assumption requires that p_{ui} should be larger than p_{uj} , $p_{ui} \succeq p_{uj}$, when user u has a direct link with user i rather than j in G . This assumption models the preference order of users when building links with others and are commonly used in one-class recommendation problems.

However, different from existing works [27, 28, 30, 19], here network topology provides us more reliable auxiliary knowledge to solve link prediction problems. Based on the point that people are more likely to make friends with the friends of their friends, we extend the original assumption to a topology-based preference assumption consisting two pairwise preference comparisons,

$$p_{ui} \succeq p_{uj}, p_{uk} \succeq p_{uj}, \quad (6) \\ \text{for } e_{ui} \text{ and } e_{ik} \in E(G), e_{uj} \text{ and } e_{uk} \notin E(G).$$

Compared this with Eq (5), here we further assume that user u prefer to make friends with k other than j since user u and i are friends and k is the friend of i .

Model Formulation. Based on the assumption, we propose an optimization criterion for each user. In particular, the inequality in Eq (6) can be used to maximize the value of the Area Under the ROC Curve (AUC) which is widely used in ranking problems. We utilize a matrix factorization framework to estimate the link preference between users. Specifically, $H \in \mathbb{R}^{M \times k}$ is a low-rank representation with $k \ll M$ for users, thus, the preference value of p_{ui} can be estimated as $H_u H_i^T + b_i$, where H_u is the u th row vector in H representing the user u 's latent factors, $b_i \in \mathbb{R}$ is a biased term for user i . Then, for each user u , our goal is to maximize the likelihood of his/her all assumed preference order as Eq (6) shows, which is equivalent to maximize Eq (7):

$$\sum_{e_{ui} \in E(G), e_{uj} \notin E(G)} \ln \mathcal{P}(p_{ui} \succeq p_{uj}) \\ + \sum_{\substack{e_{ui} \text{ and } e_{ik} \in E(G) \\ e_{uj} \text{ and } e_{uk} \notin E(G)}} \ln \mathcal{P}(p_{uk} \succeq p_{uj}). \quad (7)$$

Input: A social network $G = (U, E)$
Output: Parameters $\Theta = \{H \in \mathbb{R}^{M \times k}, b \in \mathbb{R}^M\}$
 Initialization:
 Initialize H and b
 Training:
for iterations do
 Sample user u from V uniformly;
 Sample user i and k uniformly according to
 $e_{ui}, e_{ik} \in E(G)$ and $e_{uk} \notin E(G)$;
 Calculate $h_u = H_u H^T + b$ and sort the entries in
 descending order;
 Scan from the top of h_u , find the first j satisfying
 $\Delta_u = U_u(G + e_{uh_u(v)}) - U_u(G) \leq 0$;
 for #training sample do
 Calculate $\frac{\partial \mathcal{O}(\Theta)}{\partial \Theta_{u,i,k,j}}$;
 Update H_u, H_i, H_k, H_j, b according to Eq (9);
 end
end

Algorithm 3: *Bayesian Personalized Ranking for Link recommendation with Game Theory.*

When optimizing for the AUC, it is common practice to work with a differentiable function, such as a sigmoid function which has the form $\sigma(x) = \frac{1}{1+e^{-x}}$, to approximate the function $\mathcal{P}(\cdot)$, so that the objective function is differentiable. Based on this trick, our goal is to maximize the following objective function,

$$\mathcal{O}_u(\Theta) = \sum_{u \in V} \left[\sum_{\substack{e_{ui} \in E(G) \\ e_{uj} \notin E(G)}} \ln(\sigma(p_{ui} - p_{uj})) \right. \\ \left. + \sum_{\substack{e_{ui} \text{ and } e_{ik} \in E(G) \\ e_{uj} \text{ and } e_{uk} \notin E(G)}} \ln(\sigma(p_{uk} - p_{uj})) \right] \quad (8) \\ - \text{regularization},$$

where a regularization term is used to avoid overfitting in the learning process, $p_{ui} = H_u H_i^T + b_i$, $p_{uk} = H_u H_k^T + b_k$ and $p_{uj} = H_u H_j^T + b_j$. We adopt ℓ_2 -norm regularization terms for model parameters $\Theta = \{H, b\}$.

Model Learning. We employ the widely used stochastic gradient descent (SGD) algorithm to optimize the objective function in Eq (8). The main process of SGD is to sample a training instance. Here we first randomly select a user u , then we pick a triple of users i, j, k such that $e_{ui}, e_{ik} \in E(G)$ and $e_{uk}, e_{uj} \notin E(G)$. Specifically, for each selected instance, we calculate the derivative and update the corresponding parameters Θ by walking along the ascending gradient direction,

$$\Theta^{t+1} = \Theta^t + \eta \times \frac{\partial \mathcal{O}_u(\Theta)}{\partial \Theta}. \quad (9)$$

Game Theoretic Sampling Strategy. Since the learning method we use in this paper is stochastic gradient descent (SGD), the selected training instances might have great impact on the recommendation performance. Here we adopt the utility analysis in Network Formation Game to develop a game theoretic sampling strategy for model training. In each iteration, the model needs to sample a tuple of users $\{u, i, j, k\}$. First, a user u is uniformly sampled from V , then i and k are sampled uniformly according to e_{ui} and $e_{ik} \in E(G)$, and $e_{uk} \notin E(G)$. Then we calculate a preference vector $h_u = H_u H^T + b \in \mathbb{R}^M$ for user u based on

current estimation of H . We rank the elements in h_u in descending order and let $h_u(v)$ be the user who has the v_{th} largest value in h_u . Then the game theoretic sampling strategy can be described as follows: scan from the top of h_u , find the first $h_u(v)$ satisfying $\Delta_u = U_u(G + e_{uh_u(v)}) - U_u(G) \leq 0$, set $j = h_u(v)$.

The key point of the sampling strategy is to select effective training instances by examining users' utility difference. In particular, we check the candidate j from the user with highest preference score of u , which is estimated by current model parameters. The idea is that if a user j is strongly recommended to u by current parameters but adding edge e_{uj} cannot improve u 's utility, we need to punish the model preference on j by maximizing the preference difference between p_{ui} and p_{uj} . In this way, BPRLGT not only maximize the likelihood of observed preference order but also consider the utility of each user during the learning process. The detailed algorithm is shown in Algorithm 3.

5. EXPERIMENTAL EVALUATION

In this section, we conduct several experiments to evaluate the performance of the proposed three-phase method and the unified ranking method. We demonstrate the recommendation accuracy, analyze the parameters and the convergence of the proposed methods. The results show that link recommendation performance can be improved under the help of game theoretic analysis.

5.1 Experimental Setup

First, we introduce the datasets we use in Section 3.2.2. We use four real world datasets to validate the performance of the proposed methods, namely, Ciao, Epinions, COL and Facebook. Ciao and Epinions are online product review websites. Both websites allow users to build their social connections with others in order to share their experience and knowledge on the products. COL describes a scientific collaboration network between authors submitting papers to www.arxiv.org. Facebook is one of the most popular social networking websites that users can make friends with others on it. The statistics of the four datasets are shown in Table. 1. The datasets are preprocessed by eliminating users with less than three neighbors. For each dataset, we randomly split it into a training part, used for model training, and a test part, used for model evaluation. Similar to [24], for each user u , we select 80% of their observed social links as P_u and leave the remainder as T_u for testing. Cross-validation is applied to find best parameters in the proposed methods, δ in LBR model is set as 0.25, k is set as 10 and learning rate, η , of each link prediction model is set as 0.01.

Our experiments are intended to answer the following questions:

1. Can the proposed three-phase method and unified ranking model improve the performance of existing link prediction methods?
2. Whether the proposed game theoretic sampling strategy help improve recommendation performance?
3. How does the value of k in BPRLGT affect model performance?

We use four popular metrics, $Recall@K$ ($R@K$), $NDCG$ (Normalized Discounted Cumulative Gain), Accuracy and *Area under the curve* (AUC), to measure the recommendation quality of our proposed approach in comparison to baseline methods.

The average AUC statistic is defined as

$$AUC = \frac{1}{M} \sum_{u \in M} \frac{1}{|E(u)|} \sum_{(i,j) \in E(u)} \mathbf{1}(x_{ui} > x_{uj}), \quad (10)$$

where $E(u) = \{(i, j) | (u, i) \in T_u \wedge (u, j) \notin (P_u \cup T_u)\}$ and $\mathbf{1}(\cdot)$ is an indicator function.

For each user u , $DCG@K$ considers the ranking of the recommended users by discounting the importance and is defined as

$$DCG@K = \sum_{i=1}^K \frac{2^{rel_i} - 1}{\log_2(i + 1)}, \quad (11)$$

where rel_i represents the preference score of user i (we use a binary value for this quantity, if $e_{ui} \in T_u$, $rel_i = 1$; otherwise, 0). $NDCG$ is the ratio of the DCG value to the ideal DCG value for that user. The ideal value of DCG comes from the best ranking function for the user. Here we set $K = M - |P_u|$ where M is the total number of users.

Following [32], we also use the prediction accuracy metric to measure the recommendation quality of link prediction. Specifically, the method ranks candidate links in decreasing order of confidence and we take the first $|T_u|$ candidates as the set of predicted links, denoted as C , for recommendation. Then the prediction accuracy (Accuracy) can be calculated as,

$$Accuracy = \frac{|C \cap T_u|}{|T_u|} \quad (12)$$

where $|\cdot|$ is the size of a set.

5.2 Comparison Methods

In order to demonstrate the benefits of our methods, we adopt several link prediction methods as baselines and compare their performance for link recommendation.

- **Common Neighbor (CN):** This is a basic heuristic method for link prediction which counts the number of common neighbors for link prediction.
- **Adamic-Adar (AA):** This method is an extension of common neighbor count method, where the link confidence score is computed as following,

$$AA(i, j) = \sum_{k: ik, jk \in E(G)} \frac{1}{\log \deg(k)} \quad (13)$$

where $\deg(k)$ is the degree of node k in graph.

- **BPR:** This method [27] proposes a pairwise assumption for item ranking and it is a well-known method for one-class recommendation. Here we employ it as a baseline method to compare with our proposed pairwise ranking method.
- **hTrust:** This method [32] is one of the state-of-the-art method for link prediction by exploiting homophily effect between users.
- **FactLog:** This method [24] also exploits matrix factorization framework for link prediction. Here we use log-loss with sigmoid function as the objective function with the code provided at <http://cseweb.ucsd.edu/~akmenon/code/>.
- **BPRLGT:** This is the proposed Bayesian ranking method, which extends the pairwise assumption by considering topological structure of networks and arms with a game theoretic sampling strategy.

The experimental results are shown in Table 2. In this table, for a specific method **B**, the column **B** demonstrates its performance without the help of proposed framework while **B*** describes the performance of method **B** when a NFG model is used to adjust the ranking results. The results here are achieved when LBR model is used in the framework and BPRLGT.

Table 2: Recommendation performance of different methods on four real-world datasets. For a specific method B , the column B demonstrates its performance without the help of proposed framework while B^* describes the performance of method B when a NFG model is used to adjust the ranking results. The results here are achieved when LBR model is used.

Dataset	Metrics	AA	AA*	CN	CN*	BPR	BPR*	hTrust	hTrust*	FactLog	FactLog*	BPRLGT
Ciao	R@5	0.010	0.0104	0.003	0.0037	0.007	0.0076	0.0186	0.0192	0.0275	0.0281	0.0545
	R@10	0.015	0.0159	0.011	0.01125	0.0122	0.0127	0.0375	0.0377	0.0621	0.0627	0.1044
	NDCG	0.2094	0.2099	0.1834	0.1855	0.2082	0.2090	0.2201	0.2236	0.2088	0.2092	0.3101
	Accuracy	0.081	0.086	0.032	0.034	0.134	0.148	0.172	0.179	0.159	0.163	0.1831
	AUC	0.562	0.568	0.546	0.551	0.6663	0.6703	0.6452	0.6477	0.6882	0.6884	0.9021
Epinions	R@5	0.0127	0.0129	0.0083	0.0091	0.0311	0.0323	0.0356	0.0359	0.0323	0.0330	0.040
	R@10	0.0231	0.0238	0.0162	0.0175	0.0443	0.0452	0.0547	0.0551	0.0521	0.0522	0.0620
	NDCG	0.1365	0.1373	0.124	0.1266	0.2241	0.2268	0.2378	0.2381	0.2314	0.2322	0.2485
	Accuracy	0.079	0.088	0.074	0.082	0.168	0.175	0.217	0.2202	0.1819	0.1842	0.2380
	AUC	0.5724	0.5766	0.5421	0.5508	0.8011	0.8042	0.833	0.8341	0.8211	0.8227	0.8562
COL	R@5	0.043	0.045	0.0034	0.0041	0.0012	0.0014	0.03924	0.04	0.05924	0.05930	0.0715
	R@10	0.0833	0.0841	0.0062	0.0067	0.0021	0.0025	0.08734	0.0882	0.09234	0.0928	0.1006
	NDCG	0.0853	0.0860	0.061	0.0618	0.0573	0.0582	0.1288	0.1300	0.1433	0.1444	0.1491
	Accuracy	0.0966	0.098	0.0391	0.0421	0.088	0.0894	0.118	0.121	0.1392	0.1397	0.1481
	AUC	0.761	0.7682	0.6233	0.6313	0.6174	0.6188	0.7323	0.7355	0.8232	0.8244	0.8440
Facebook	R@5	0.0032	0.00326	0.0017	0.00174	0.0025	0.0256	0.0275	0.0278	0.0321	0.0323	0.0370
	R@10	0.0048	0.0051	0.00246	0.00251	0.0044	0.0048	0.0532	0.0538	0.055	0.0562	0.0618
	NDCG	0.2214	0.2218	0.1832	0.1838	0.2191	0.2199	0.2934	0.2945	0.3341	0.3352	0.3740
	Accuracy	0.0126	0.0132	0.0042	0.0051	0.009	0.0093	0.0623	0.0643	0.0681	0.0683	0.0769
	AUC	0.6682	0.6685	0.5823	0.5848	0.6461	0.6488	0.7245	0.7248	0.7823	0.7844	0.9642

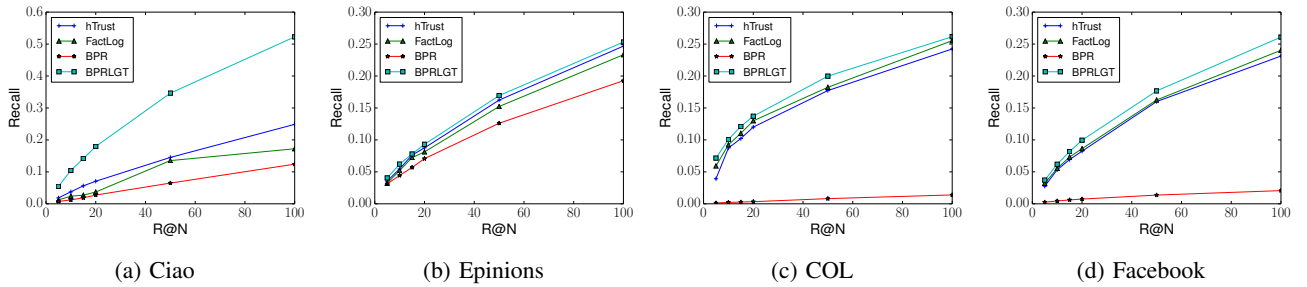


Figure 2: Recommendation performance comparisons (Recall@N) on four datasets.

5.3 Summary of Experimental Results

Performance Comparison. Based on the results in Table 2, we can see that the proposed unified model clearly outperforms the comparison methods. CN and AA are two static scoring methods using only local network structure information for link prediction, which often lead to suboptimal performance. BPR is a popular pairwise ranking method for one-class recommendation problem and here we use it to model users' preference order when building links. However, it does not perform well because it misses network topology and just uses a uniform sampling strategy for model training. $hTrust$ is a state-of-the-art point wise method for link prediction, incorporating low-rank matrix factorization and homophily regularization terms. $FactLog$ also exploits matrix factorization framework for link prediction and also take advantage of network topology as explicit feature for model learning. Although both $hTrust$ and $FactLog$ methods improve the performance of link prediction compared with BPR and other heuristic methods, they still cannot beat the proposed $BPRLGT$ model. $BPRLGT$ carries on the merits of BPR and further extends BPR 's assumption by directly modeling users' preference on their second-degree friends. Moreover, when learning the model, the proposed game theoretic sampling strategy allows $BPRLGT$ to consider users' utility difference for effective

training instance selection. Based on these, $BPRLGT$ obtains significant improvements over all baseline methods on all datasets.

Detailed results for $Recall@N$ can be found in Figure 2. We can see that on Ciao dataset, $BPRLGT$ achieves clearly better performance than other methods. For other datasets, although the curves of different methods seem to be close to each other, $BPRLGT$ consistently outperform baselines on different $Recall@N$ measure points. We also find that compared with $BPRLGT$, BPR always performs bad, especially in COL dataset. The reason we consider is that BPR samples training instances via uniform sampling and do not consider network topology information. Since users' observed links are always very limited and there are huge number of unobserved links acting as candidate links to be recommended, thus the original assumption of BPR , Eq (5), cannot correctly capture users' preference on building links. This also indicates the necessity of proposing new pairwise preference assumption for link prediction, as $BPRLGT$ does.

We also take experiments to evaluate the performance of the proposed framework. For a specific method B , the column B demonstrates its performance without the help of proposed framework while B^* describes the performance of method B when a NFG model is used to adjust the ranking results. The results here are achieved when LBR model is used since LBR performs better in

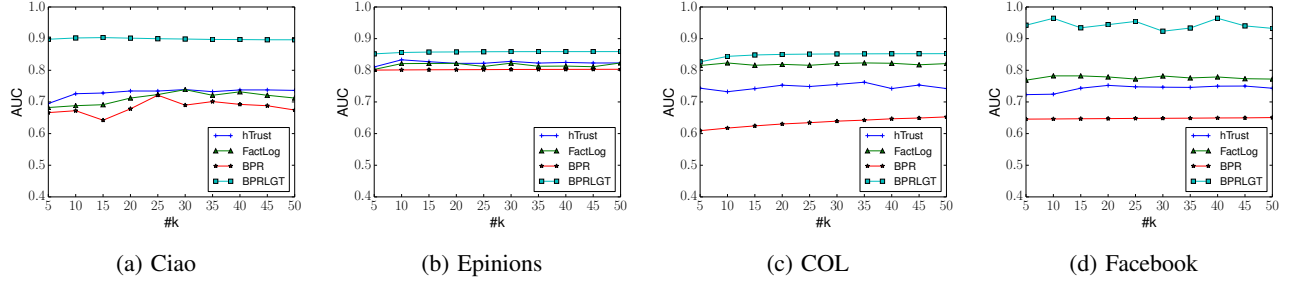


Figure 3: Effects of latent factor number.

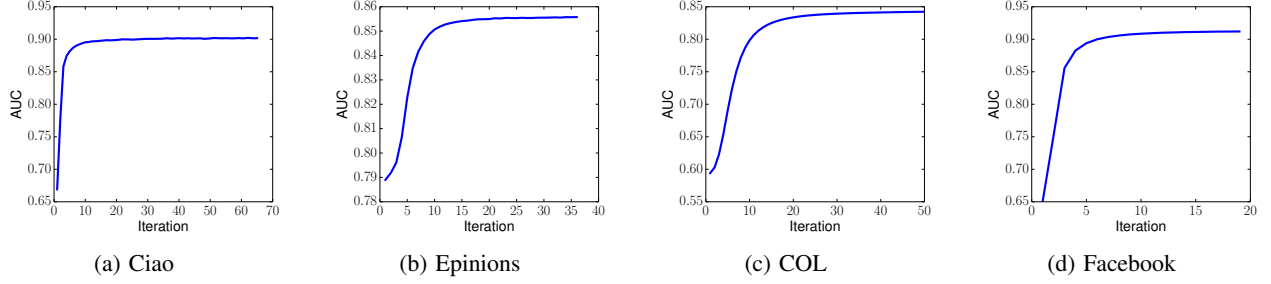


Figure 4: Convergence analysis.

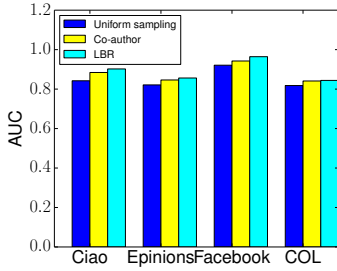


Figure 5: Effects of sampling strategies.

Sec 3.2.2. From the results, on all datasets, we find that exploiting game theoretic analysis can improve prediction performance for all link prediction methods. By exchanging the positions of lower-ranked but beneficial links with those higher-ranked but non-beneficial links, our proposed framework can help link prediction methods achieve better performance, which verifies the contribution of Network Formation Game. In summary, with the help of game theoretic analysis, both the proposed framework and the unified ranking model can always improve the performance of baseline methods.

5.4 Analysis and Discussion

Impact of k . Since *BPRLGT* is built upon matrix factorization framework, which requires us to predefine the number of latent factors in advance, here we take experiments to investigate the impact of the number of latent factor k on recommendation performance. The results are shown in Figure 3. We use AUC as evaluation metric and vary the value of k from 5 to 50. We find that on all datasets, all comparison methods perform stable with different values of k

and *BPRLGT* always achieves best and stable performance. The results verify the significant contribution of the extended preference assumption as Eq (6) and game theoretic sampling strategy. Moreover, we also notice that on COL and Facebook datasets, *BPR* shows clearly poor performance than other methods. The reasons we consider are that the observed feedback in these datasets is too sparse and *BPR* does not consider network topology to help select effective instances for better learning.

Impact of Game Theoretic Sampling Strategy. In this subsection, we compare the performance when we use different Network Formation Game models in *BPRLGT*. The results are shown in Figure 5. We find that on all datasets, *BPRLGT* with game theoretic sampling strategy achieves better performance over uniformly sampling strategy. Moreover, by using *LBR* model, it achieves best performance over Co-author models. Through cross validation, δ and \bar{x} here are set as 0.25 and 1 in *LBR*, respectively. This results verify the simulation results in Section 3.2.2 and also demonstrate the advantages of incorporating game theoretic analysis into machine learning models for link prediction.

Convergence Analysis. We further investigate the convergence of *BPRLGT*. Figure 4 shows the convergence analysis of *BPRLGT* model on different datasets. We see that our method always converges within 30 iterations on all datasets, since *BPRLGT* incorporates network topology into the model assumption and also benefits from utility analysis from NFG models, making it learn more effective knowledge in each iteration. The fast convergence ensures its ability to scale up to huge data.

6. CONCLUSION

In this paper, we study the problem of exploiting game theoretic analysis for link recommendation in social networks. We first propose a three-phase game theoretic framework combining the models in Network Formation Game with the machine learning models.

Besides this framework, we also develop a unified model by incorporating Network Formation Game into a Bayesian ranking model for link recommendation. Specifically, we design a game theoretic sampling approach to improve the model training process. Experiments on four real datasets validate the effectiveness of the proposed methods for link recommendation.

The idea and methods proposed in this paper represents a novel and interesting research direction. There are several potential directions for future work. One straightforward extending is to study the performance when other Network Formation models are incorporated into the proposed framework and ranking model. Moreover, it is also interesting to investigate how to combine contextual knowledge with Network Formation models for link recommendation. Furthermore, integrating game theory with other machine learning models, such as Markov Decision Process, should also be an intriguing direction for future research.

7. ACKNOWLEDGEMENTS

This research was in part supported by grants from the National Grand Fundamental Research 973 Program of China (No. 2014CB340405), the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 413213), and Microsoft Research Asia Regional Seed Fund in Big Data Research (Grant No. FY13-RES-SPONSOR-036).

8. REFERENCES

- [1] R. Andersen, C. Borgs, J. Chayes, U. Feige, A. Flaxman, A. Kalai, V. Mirrokni, and M. Tennenholtz. Trust-based recommendation systems: an axiomatic approach. *In WWW*, 2008.
- [2] A. I. Badev. Social networks and the stepping stone effect: The case of tobacco and marijuana. *working paper*, 2014.
- [3] Ballester, Coralio, and A. CalvU-Armengol. Moderate interactions in games with induced complementarities. 2007.
- [4] Ballester, Coralio, A. CalvU-Armengol, and Y. Zenou. Who's who in networks. wanted: the key player. *Econometrica*, 2006.
- [5] M. Bayati, C. Borgs, J. Chayes, Y. Kanoria, and A. Montanari. Bargaining dynamics in exchange networks. 2010.
- [6] G. Beigi, H. Alvari, M. Jalili, and G. Sukthankar. Leveraging community detection for accurate trust prediction. *ASE ICSC*, 2014.
- [7] Y. Bramoulle, R. Kranton, and M. DiAmours. Strategic interaction and networks. *CIRPEE Working Paper*, 2011.
- [8] S. Currarini and M. Morelli. Network formation with sequential demands. *Review of Economic Design*, 2010.
- [9] L.-N. David and K. Jon. The link prediction problem for social networks. *CIKM*, 2003.
- [10] Y. Dong, J. Tang, S. Wu, J. Tian, N. Chawla, J. Rao, and H. Cao. Link prediction and recommendation across heterogenous social networks. *In ICDM*, 2012.
- [11] B. Dutta, S. Ghosal, and D. Ray. Farsighted network formation. *Journal of Economic Theory*, 2005.
- [12] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. *In WWW*, 2004.
- [13] Y. Hu, Y. Koren, and C. Volin-sky. Collaborative filtering for implicit feedback datasets. *In ICDM*, pages 263–272, 2008.
- [14] M. Jackson and Y. Zenou. Games on networks. *Handbook of Game Theory*, 2012.
- [15] M. O. Jackson. A survey of models of network formation: Stability and efficiency. *Group Formation in Economics. Networks, Clubs and Coalitions*, 2004.
- [16] M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, pages 44–74, 1996.
- [17] P. Kohli, Y. Bachrach, T. Graepel, G. Smyth, M. Armstrong, D. Stillwell, and M. Kearns. Behavioral game theory on online social networks: Colonel blotto is on facebook. 2011.
- [18] P. Kohli, Y. Bachrach, D. Stillwell, M. Kearns, and R. Herbrich. Colonel blotto on facebook: The effect of social relations on strategic interaction. *Proceedings of the 3rd Annual ACM Web Science Conference*, 2012.
- [19] A. Krohn-Grimberghe, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme. Multi-relational matrix factorization using bayesian personalized ranking for social network data. *In WSDM*, pages 173–182, 2012.
- [20] C. Lee, M. Pham, N. Kim, M. K. Jeong, D. K. Lin, and W. A. Chavalitwongse. A novel link prediction approach for scale free networks. *In WWW*, 2014.
- [21] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. *In WWW*, 2010.
- [22] W. S. Lin, H. V. Zhao, and K.J.RayLiu. Game-theoretic strategies and equilibriums in multimedia fingerprinting social networks. *IEEE TRANSACTIONS ON MULTIMEDIA*, 2011.
- [23] A. Mele. A structural model of segregation in social networks. *working paper*, 2010.
- [24] A. K. Menon and C. Elkan. Link prediction via matrix factorization. *In ECML/PKDD*, 2011.
- [25] R. Pan, Y. Zhou, B. Cao, N. N. Liu, R. Lukose, M. Scholz, and Q. Yang. One-class collaborative filtering. *In ICDM*, pages 502–511, 2008.
- [26] A. D. Paula, S. Richards-Shubik, and E. Tamer. Identification of preferences in network formation games. 2015.
- [27] S. Rendle, C. Freuden-thaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: bayesian personalized ranking from implicit feedback. *In UAI*, pages 452–461, 2009.
- [28] S. Rendle, C. Freuden-thaler, and L. S. Thieme. Factorizing personalized markov chains for next-basket recommendation. *In WWW*, pages 811–820, 2010.
- [29] S. Rendle and C. Freudenthaler. Improving pairwise learning for item recommendation from implicit feedback. *In WSDM*, 2014.
- [30] S. Rendle and L. Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. *In WSDM*, pages 81–90, 2010.
- [31] B. Skyrms and Pemantle. A dynamic model of social network formation. *National Academy of Sciences*, 2000.
- [32] J. Tang, H. Gao, X. Hu, and H. Liu. Exploiting homophily effect for trust prediction. *In WSDM*, 2013.
- [33] J. Tang, H. Gao, and H. Liu. mtrust: Discerning multi-faceted trust in a connected world. *In WSDM*, 2012.
- [34] J. Tang, H. Gao, H. Liu, and A. D. Sarma. etrust: Understanding trust evolution in an online world. *In KDD*, 2012.
- [35] L. Xiao, Y. Chen, W. S. Lin, and K. J. R. Liu. Indirect reciprocity security game for large-scale wireless networks. *IEEE Transactions on Information Forensics and Security*, 2012.
- [36] Y. Yao, H. Tong, X. Yan, F. Xu, and J. Lu. Matri a multi-aspect and transitive trust inference model. *In WWW*, 2013.
- [37] G. Zappella, A. Karatzoglou, and L. Baltrunas. Games of friends: A game-theoretical approach for link prediction in online social networks. *Workshop in AAAI*, 2013.
- [38] T. Zhao, J. J. McAuley, and I. King. Leveraging social connections to improve personalized ranking for collaborative filtering. *In In CIKM*, 2014.
- [39] C. Ziegler and J. Golbeck. Investigating interactions of trust and interest similarity. *Decision Support Systems*, 2007.