

Opinions Power Opinions: Joint Link and Interaction Polarity Predictions in Signed Networks

Tyler Derr¹, Zhiwei Wang¹, and Jiliang Tang¹

Abstract—Social media has been widely adopted by online users to share their opinions. Among users in signed networks, two types of opinions can be expressed. They can directly specify opinions to others via establishing positive or negative links; and they also can give opinions to content generated by others via a variety of social interactions such as commenting and rating. Intuitively these two types of opinions should be related. For example, users are likely to give positive (or negative) opinions to content from those with positive (or negative) links; and users tend to create positive (or negative) links with those that they frequently positively (or negatively) interact with. Therefore we can leverage one type of opinions to power the other. Meanwhile, they can enrich each other that can help mitigate the data sparsity and cold-start problems in the corresponding predictive tasks – link and interaction polarity predictions, respectively. In this paper, we investigate the problem of joint link and interaction polarity predictions in signed networks. We first understand the correlation between these two types of opinions; and then propose a framework that can predict signed links and the polarities of interactions simultaneously. The experimental results on a real-world signed network demonstrate the effectiveness of the proposed framework. Further experiments are conducted to validate the robustness of the proposed framework to data with cold-start users.

I. INTRODUCTION

Traditionally, network analysis has focused on unsigned networks (or networks with only positive links). Many social networks in social media can have positive and negative links (or signed networks [1], [2]). Such examples include the Epinions network with trust and distrust and the Slashdot network with friend and foe links. Negative links have been proven to advance various network analysis tasks such as link prediction [3], [4], [5], node classification [6], community detection [7], [8], [9], and recommendations [10], [11], [12]. Meanwhile, a recent study has shown that invisible negative links in social media are predictable [13] and that they can help convert many social media unsigned networks such as Facebook friendship and Twitter following into signed networks. Therefore, signed networks are ubiquitous and have attracted increasing attention in recent years [14].

Meanwhile, social media has been increasingly used by online users to share and exchange opinions. In signed networks, users can directly express positive (or negative) opinions to others by establishing positive (or negative) links. They can also specify positive (or negative) opinions to content created by others via various interactions such as commenting and rating. These two types of opinions should

be related inherently. For example, a user receiving more positive (or negative) links is likely to receive more positive (or negative) opinions for his/her content; while users are likely to give positive (or negative) opinions to content generated by those with positive (or negative) links. In reality, users may also explicitly only give opinions to a small number of users or content. For example, both positive and negative links follow a power-law-like distribution – a small number of users specify many positive (or negative) links while a large proportion of users specify a few positive (or negative) links [15]. Hence, link prediction [5] and interaction polarity prediction [16] are proposed to infer implicit opinions of these two types, respectively.

Recent years have witnessed a large number of algorithms for signed link prediction [5], [17], [4], [18] and interaction polarity prediction [19], [20], [16], [21]. However, the majority of them have tackled these two tasks independently. As aforementioned, the corresponding opinions in these two tasks could be correlated and we can utilize one to power the other. Thus, we could boost the performance by jointing these two tasks. Meanwhile, due to the sparsity nature of social media data, both tasks have been shown to severely suffer from the data sparsity and cold-start problems [7], [21]. By capturing the correlation between these two types of opinions, one can enrich the other and therefore have more information to use for their corresponding tasks. Hence, a joint framework has the potential to mitigate the data sparsity and cold-start problems for both tasks.

In this paper, we study the problem of joint link and interaction polarity predictions in signed networks. In particular, we investigate – (a) whether opinions in the two tasks are related? and (b) how to utilize their correlations for joint link and interaction polarity predictions? Providing answers to these two questions, we propose a novel framework LIP that can infer links and polarities of interactions jointly. Our main contributions in this work have been summarized as follows:

- We validate the correlations between link signs and interaction polarities from both global and local perspectives;
- We propose a joint link and interaction prediction framework (LIP) that explicitly incorporates the correlations to predict links and interaction polarities simultaneously;
- We conduct experiments in a real-world signed network to demonstrate (a) the effectiveness of LIP and (b) the robustness of LIP to the data sparsity and cold-start problems.

¹T. Derr, Z. Wang, and J. Tang are with the Data Science and Engineering Lab, Michigan State University, East Lansing, MI 48824, USA {derrtyl, wangzh65, tangjili} at msu.edu

The rest of this paper is organized as follows. In Section 2, we formally define the joint prediction problem. In Section 3, we describe the dataset used in our work, along with our analysis of the correlations. We discuss our proposed novel joint framework in Section 4. In Section 5 the experimental results and findings are presented. We briefly review related work in Section 6. Conclusions and future work are given in Section 7.

II. PROBLEM

Let $\mathcal{U} = \{u_1, u_2, \dots, u_n\}$ denote the set of n users. We represent signed links between users in an adjacency matrix, $\mathbf{T} \in \mathbb{R}^{n \times n}$, where $\mathbf{T}_{ij} = 1$ if u_i creates a positive link to u_j , -1 if u_i creates a negative link to u_j , and 0 otherwise (i.e., when u_i has shown no link to u_j). Let $\mathcal{R} = \{r_1, r_2, \dots, r_m\}$ be the set of m content items generated by \mathcal{U} . We use $\mathbf{A} \in \mathbb{R}^{n \times m}$ to denote the authorship matrix where $\mathbf{A}_{ij} = 1$ if u_i creates r_j and $\mathbf{A}_{ij} = 0$ otherwise. Social media provides multiple ways for users to express their opinions to content items generate by other users. For example, Facebook and Twitter allow their users to comment on content; Youtube provides thumbs-up and -down buttons; and Epinions enables its users to rate the helpfulness of the content with scores from 1 to 6. We use $\mathbf{H} \in \mathbb{R}^{n \times m}$ to denote opinions expressed by \mathcal{U} to \mathcal{R} , where $\mathbf{H}_{ik} = 1$ (or -1) if u_i gives a positive (or negative) opinion to r_k and we use $\mathbf{H}_{ik} = 0$ to indicate no explicit opinion is expressed from u_i to r_k . Note that in this paper, we define positive (or negative) interactions between u_i and u_j as u_i giving positive (or negative) opinions to content items generated by u_j . In other words, an interaction between users is defined as a triplet (u_i, r_k, u_j) where u_i expresses opinions to r_k that was generated by u_j .

With the above notations and definitions, our problem is stated as follows: *given the signed relations \mathbf{T} , the authorship matrix \mathbf{A} and the user-item opinion matrix \mathbf{H} , we aim to learn a predictor that can infer signed links and interaction polarities simultaneously by leveraging \mathbf{T} , \mathbf{A} , and \mathbf{H} .*

Note that when the content of the item is available, we also can utilize the content of \mathcal{R} . However, in this paper, we focus on leveraging \mathbf{T} , \mathbf{A} , and \mathbf{H} and would like to leave the problem of exploiting content as one future work.

III. DATA ANALYSIS

In this section, we conduct preliminary analysis on the correlation between signed links and interaction polarities. We begin by introducing the dataset in our study.

A. Dataset

We collected a dataset from Epinions for this investigation. Epinions users can give positive and negative links to each other, which we use to construct the \mathbf{T} matrix. They also can write reviews and we use this data to construct the authorship matrix \mathbf{A} . For each review others can use scores from 1 to 6 to indicate the helpfulness of the given reviews and that we use these to construct the matrix \mathbf{H} . We define positive and negative helpfulness ratings to be $\{4, 5, 6\}$ and $\{1, 2, 3\}$, respectively. Some statistics of the dataset are shown in Table I. From the table, we can observe that (1) there are

TABLE I
EPINIONS DATASET STATISTICS.

# of Users	233,429
# of Positive Links	717,667
# of Negative Links	123,705
Density of \mathbf{T}	7.75×10^{-5}
# of Reviews	755,722
# of Positive Interactions	12,581,553
# of Negative Interactions	1,086,551
Density of \mathbf{H}	1.54×10^{-5}

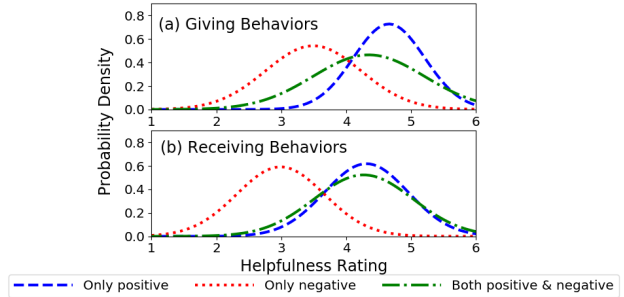


Fig. 1. A Global Perspective on Opinion Correlations.

more positive links (or interactions) than negative ones; and (2) both links and interactions are very sparse. The task of creating (or receiving) a signed link to others can be thought of as an explicit form of expressing one's opinion of (or from) others. In contrast, when a user interacts with the content authored by others, they are implicitly marking their opinion towards others in these interactions. Therefore, it is reasonable to assume that the implicit and explicit opinions among users are correlated. Next we investigate these correlations from both global and local perspectives.

B. A Global Perspective

From a global perspective, we want to examine the correlations between these explicit and implicit opinions from one user. In particular, we aim to answer the following questions – (1) is a user, giving more positive (or negative) links, likely to give more positively (or negatively) on content from others? and (2) is a user, receiving more positive (or negative) links, likely to receive more positive (or negative) opinions on his/her content? In this work, we refer to giving links or opinions on content as giving behaviors; while receiving links or opinions on content as receiving behaviors.

To answer the first question, we group users into three classes based upon their outgoing links as follows: (1) users who only have positive outgoing links (76,819 users); (2) users having only negative outgoing links (7,138 users); and (3) users who have both positive and negative outgoing links (11,361 users). Then, we calculate the opinions (or helpfulness ratings) they gave to content from others for each group and we plot kernel smoothing density estimation for each group in Figure 1(a). We note that on average users who only create positive links also tend to interact more positively with the content generated by other users as compared to users who only create negative links. Furthermore, the users who create both positive and negative links are shown to express both positive and negative behaviors when examining their interactions performed on content generated by other

users. Evidence from the figure supports that users who give more positive (or negative) links tend to express positively (or negatively) on content from others.

To answer the second question, we divide users into three groups based upon their incoming links as follows: (1) users who only have positive incoming links (52,810 users); (2) users having only negative incoming links (14,701 users); and (3) users who have both positive and negative incoming links (17,090 users). Following the similar procedure, we plot kernel smoothing density estimation of receiving behaviors for each group in Figure 1(b). From the figure, we can make very similar observations for receiving behaviors as giving behaviors, which lead to a positive answer to the second question – users, receiving more positive (or negative) links, are likely to obtain more positive (or negative) opinions on their content.

C. A Local Perspective

The global perspective in Subsection III-B focuses on correlations between one user and the remaining network. In this subsection, we focus on a pair of users and we want to investigate whether the existence of a positive (or negative) link for a pair of users makes a difference on how they give (or receive) opinions on each other’s content. In particular, for a pair of users u_i to u_j , we aim to answer – (1) if u_i gives a positive (or negative) link to u_j , is u_i likely to give positive (or negative) opinions to content from u_j ? ; and (2) if u_j receives a positive (or negative) link from u_i , is u_j likely to give positive (or negative) opinions to the content from u_i ? Note that in this work, we use $u_i + u_j$, $u_i - u_j$ and $u_i ? u_j$ to denote a positive, negative and no link from u_i to u_j .

To answer the first question, we divide all pairs of users into three groups – (a) positive pairs $u_i + u_j$; (b) negative pairs $u_i - u_j$; and no-link pairs $u_i ? u_j$. For each pair in each group, we calculate the average opinion (or helpfulness ratings) from u_i to the content of u_j . We apply kernel smoothing density estimation for each group and the distributions are shown in Figure 2(a). From the figure, we note that on average positive pairs have higher helpfulness scores than no-link pairs, which have higher scores than negative pairs. Hence, it is quite evident from the figure that if u_i gives a positive (or negative) link to u_j , u_i is likely to give positive (or negative) opinions to the content from u_j .

Intuitively, if u_j receives a positive link from u_i , u_j is likely to be friendly to u_i , and as a consequence, u_j is likely to give positive opinions to the content of u_i . On the other hand, if u_j receives a negative link from u_i , u_j could do revenge back and give negative opinions to the content of u_i . We follow a similar procedure of answering the first question for the second question. The results are demonstrated in Figure 2. From the figure, we observe that (1) on average, u_j mostly gives positive opinions to the content from those who give positive links to u_j ; while u_j mostly gives negative opinions to the content from those who give negative links to u_j . These observations support that if u_j receives a positive

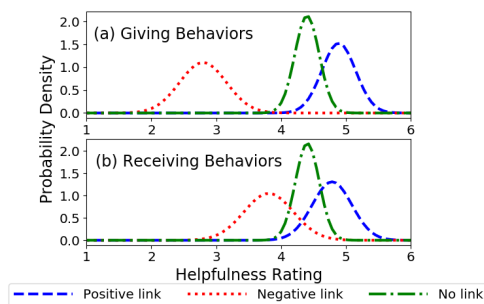


Fig. 2. A Local Perspective on Opinion Correlations.

(or negative) link from u_i , then u_j is likely to give opinions being more positive (or negative) to the content from u_i .

IV. A FRAMEWORK FOR JOINT LINK AND INTERACTION POLARITY PREDICTIONS

In Section III, we validated that there exist correlations between a user’s opinion of other users in regards to the links they form in signed social networks and the polarities of the interactions between them. Thus, these findings naturally lead us to the question of whether this knowledge can benefit the two prediction tasks that are found in the two domains; link and interaction polarity prediction. In this section, we first briefly discuss a basic framework to solve the two tasks of link and interaction polarity predictions individually. Then we discuss how to model the opinion correlations that enable us to have the opinions in one task power the other. Finally we present our proposed framework LIP, which directly incorporates these correlations into a joint optimization algorithm that can infer links and polarities of interactions jointly.

A. Basic Prediction Models

The low-rank matrix factorization approach has gained popularity recently and is now being used across various applications such as link prediction [22], [4] and recommender systems [23], [21]. In this work, we choose to build the basic prediction models based on the low-rank matrix factorization approach.

1) *Link Prediction*: Let $\mathcal{T} = \{(u_i, u_j) | \mathbf{T}_{ij} \neq 0\}$ be the set of pairs with links. In terms of the link prediction task, we would like to find two latent matrices $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in \mathbb{R}^{K_L \times n}$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{R}^{K_L \times n}$, with K_L being the number of latent dimensions, by solving the following optimization problem:

$$\min_{\mathbf{U}, \mathbf{V}} \frac{1}{2} \sum_{(u_i, u_j) \in \mathcal{T}} (\mathbf{T}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2 + \frac{\beta_1}{2} (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2) \quad (1)$$

where \mathbf{u}_i and \mathbf{v}_i are the user latent vectors representing giving and receiving link behaviors of u_i , respectively. Thus, $\mathbf{u}_i^\top \mathbf{v}_j$ models the sign of a link from u_i to u_j , and therefore after optimizing the above formulation, we can use such inner products as a prediction for unknown user-user signed links in the network. Note that $\|\mathbf{U}\|_F^2$ denotes the Frobenius norm of \mathbf{U} and is used as a regularization term to prevent overfitting, similarly for \mathbf{V} , and both are controlled by the parameter β_1 .

2) *Interaction Polarity Prediction*: Let $\mathcal{H} = \{(u_i, r_k, u_j) | \mathbf{H}_{ik} \neq 0, \mathbf{A}_{jk} \neq 0\}$ be the set of interaction triplets and \mathbf{H}_{ik} denotes the opinion from u_i to the content r_k authored by u_j . The main difference between the basic model for this task from traditional matrix factorization based recommender systems is that we now have a third piece of information, the author. Thus, rather than taking the typical user-item formulation, we instead want to formulate the model so that we can include information about the author of the content.

In this problem, we wish to find three latent matrices $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n] \in \mathbb{R}^{K_I \times n}$, $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n] \in \mathbb{R}^{K_I \times n}$, and $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m] \in \mathbb{R}^{K_I \times m}$, where \mathbf{p}_i and \mathbf{q}_i respectively denote the giving and receiving interaction behaviors of u_i , and \mathbf{s}_k is the latent vector for content r_k . These three matrices can be obtained via solving the following optimization problem:

$$\min_{\mathbf{P}, \mathbf{Q}, \mathbf{S}} \frac{1}{2} \sum_{(u_i, r_k, u_j) \in \mathcal{H}} (\mathbf{H}_{ik} - \mathbf{p}_i^\top (\mathbf{q}_j + \mathbf{s}_k))^2 \quad (2)$$

$$+ \frac{\beta_2}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2 + \|\mathbf{S}\|_F^2)$$

the term $(\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2 + \|\mathbf{S}\|_F^2)$ is introduced to avoid overfitting, which is controlled by β_2 . Next we will discuss how to capture correlations based on the two aforementioned basic models.

B. Modeling Opinion Correlations

In Section III, we found that the giving (or receiving) behaviors in terms of links and interactions are correlated. In the basic models from Subsection IV-A.2, we use \mathbf{u}_i and \mathbf{v}_i to denote users' behaviors when giving and receiving links, respectively. While we use \mathbf{p}_i and \mathbf{q}_i to respectively indicate users' behaviors when giving and receiving interactions, separately. Therefore, we can capture the opinion correlations by bridging the two giving behaviors via \mathbf{u}_i and \mathbf{p}_i , and the two receiving behaviors via \mathbf{v}_i and \mathbf{q}_i . Since the two giving behaviors are correlated, we can find a linear mapping matrix $\mathbf{W}_O \in \mathbb{R}^{K_I \times K_L}$ that can map u_i 's latent vector \mathbf{u}_i , which denotes his/her underlying behavior on how to create links, to the latent vector \mathbf{p}_i , which captures their behavior towards how they give opinions to the content authored by other users in the network. Given a set of latent vectors for all users $u_i \in \mathcal{U}$, it can then be easily seen that the linear mapping between them would be a solution to the following optimization problem:

$$\min_{\mathbf{W}_O} \sum_{u_i \in \mathcal{U}} \|\mathbf{W}_O \mathbf{u}_i - \mathbf{p}_i\|_2^2 \quad (3)$$

Similarly, we seek to find a matrix $\mathbf{W}_I \in \mathbb{R}^{K_I \times K_L}$ to represent the mapping between the user u_j 's latent vectors \mathbf{v}_j , and \mathbf{q}_j , which denote their receiving behaviors of receiving links and interactions, respectively. The mapping \mathbf{W}_I can be learned as follows:

$$\min_{\mathbf{W}_I} \sum_{u_j \in \mathcal{U}} \|\mathbf{W}_I \mathbf{v}_j - \mathbf{q}_j\|_2^2 \quad (4)$$

Eqs. (3) and (4) can capture opinion correlations for links and interactions. They also allow us to bridge the two basic models for link and interaction polarity predictions together. Next we will introduce the proposed joint framework.

C. The Proposed Joint Framework

Now we have formulated a model on how to optimize a linear mapping between both the giving and receiving behaviors in the two tasks. Next we show how these mappings can be used as two additional terms in our joint matrix factorization framework, LIP, for the purpose of joint link

and interaction polarity prediction. LIP solves the following optimization problem:

$$\min_{\substack{\mathbf{U}, \mathbf{V}, \mathbf{P}, \mathbf{Q}, \\ \mathbf{S}, \mathbf{W}_I, \mathbf{W}_O}} \mathcal{L}(\mathbf{U}, \mathbf{V}, \mathbf{P}, \mathbf{Q}, \mathbf{S}, \mathbf{W}_I, \mathbf{W}_O)$$

$$= \frac{1}{2} \sum_{(u_i, u_j) \in \mathcal{T}} (\mathbf{T}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2$$

$$+ \frac{\eta}{2} \sum_{(u_i, r_k, u_j) \in \mathcal{H}} (\mathbf{H}_{ik} - \mathbf{p}_i^\top (\mathbf{q}_j + \mathbf{s}_k))^2$$

$$+ \frac{\gamma}{2} \left(\sum_{u_i \in \mathcal{U}} \|\mathbf{W}_O \mathbf{u}_i - \mathbf{p}_i\|_2^2 + \sum_{u_j \in \mathcal{U}} \|\mathbf{W}_I \mathbf{v}_j - \mathbf{q}_j\|_2^2 \right)$$

$$+ \frac{\beta_1}{2} (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2) + \frac{\beta_2}{2} (\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2 + \|\mathbf{S}\|_F^2)$$

$$+ \frac{\beta_3}{2} (\|\mathbf{W}_I\|_F^2 + \|\mathbf{W}_O\|_F^2) \quad (5)$$

where the first term is a standard user-user matrix factorization model (as discussed in Subsection IV-A) for the link prediction problem. The second term is a modification to the user-review matrix factorization model that also incorporates the additional vector $\mathbf{q}_j \forall u_j \in \mathcal{U}$ to represent the influence of the author u_j in the prediction of u_i 's opinion on r_k , when r_k was written by u_j . The third and fourth terms capture the correlations of giving and receiving behaviors, respectively, and their contributions are controlled by a parameter γ . Other terms in Eq. (5) are added to avoid overfitting.

We note that the balance between optimizing for the two tasks (sign link prediction and user interactions polarities) is balanced by the parameter η , where a small increase in this value will result in an increase to the importance of the user interaction polarity prediction task, and similarly towards the link prediction task when decreasing its value. Also, this transfer of information between problems is done by the linear mapping used in LIP (more specifically the terms controlled by γ in Eq. (5)). If a user u_i has no link information, they are deemed a cold-start user in the link prediction task. Thus there is no way to learn \mathbf{u}_i and \mathbf{v}_i in the basic model and we fail to do link prediction for u_i . However, if u_i has had some interactions with other users in the network, we can learn \mathbf{p}_i and \mathbf{q}_i from his/her interaction data. Thus, the proposed framework LIP can also learn \mathbf{u}_i and \mathbf{v}_i via the model components of capturing giving and receiving correlations via the third and fourth terms in Eq. (5). Similarly, LIP can also help when u_i has no interaction data but has link information. Via the above analysis, we note that LIP has the potential to mitigate the data sparsity and cold-start problems in either link prediction or interaction polarity prediction.

D. An Optimization Method for LIP

Given the the optimization objective shown above, we now present how to solve this problem. We have chosen to use stochastic gradient descent (SGD) due to the non-convexity of the joint optimization formulation. First, we compute the partial derivatives with respect to each of the parameters (i.e., $\mathbf{u}_i, \mathbf{v}_j, \mathbf{p}_i, \mathbf{q}_j, \mathbf{s}_k, \mathbf{W}_O$, and \mathbf{W}_I) and then iteratively update them using SGD until convergence. We use the combined training data $\mathcal{X} = \{\mathcal{T} \cup \mathcal{H}\}$, where \mathcal{T} and \mathcal{H} are the link and interaction training data, respectively.

V. EXPERIMENTS

In this section, we conduct experiments to answer the following two questions: (1) Can our joint model help alleviate the sparsity problem in these two prediction tasks? (2) Do the terms based upon correlated user opinions/behaviors in LIP provide a transfer of information between the two problems? To address the first question, we perform experiments in which we increase the sparsity of the training data and compare the performance with representative baselines. We address the second question by examining if our algorithm is robust to handle some cold-start users. In the next subsection we will further introduce our dataset and how it was used, the metric used in evaluating the two prediction tasks, then we introduce the experimental settings for the two types of experiments we have performed.

A. Experimental Settings

As mentioned in Section III, we have collected a dataset from Epinions for these experiments. Note that for the purpose of this study, we have filtered our collected Epinions dataset to form more dense user-user and user-content matrices. The first step is to pre-process the data such that we have the appropriate training, validation, and testing sets from our dataset.

The filtering we perform only keeps users that have both given and received a link, and also requires the users to have given at least one helpfulness rating and have also authored at least one review that has received at least one helpfulness rating. For all selected users to be filtered out, we remove all their user links, reviews they had written, and helpfulness ratings associated with that user. The reason for this filtering is that it will allow us to later remove portions of the data to artificially create training sets that have a varying percentage of cold start users and also different levels of sparsity.

The original dataset had contained 233,429 users, 841,373 user-user links, and 13,668,105 helpfulness ratings. After the above mentioned filtering process, we were left with 29,901 users, 600,976 user-user links, and 11,555,599 helpfulness ratings. The dataset has been randomly split into 70% for training, 10% for validation, and 20% for testing. Note that we then balanced our testing dataset to be 50% positive and 50% negative similar to that done in [5]. To evaluate and compare the performance of LIP, we present the F1 measure for the interaction polarity and the link prediction tasks. Note that the higher the value, the better the performance.

For all the models that required parameters to be tuned, we used the validation set to obtain the best parameters for each respective model.

B. Sparsity Experiments

To answer the first question, we compare the proposed framework, LIP, with existing interaction polarity and link prediction methods. We first present the baselines for the interaction polarity prediction task followed by those for the link prediction task.

We choose the following representative interaction polarity prediction baselines for comparison:

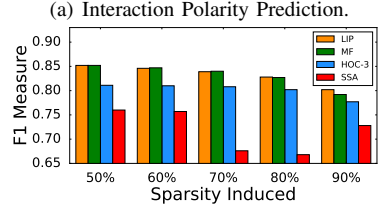
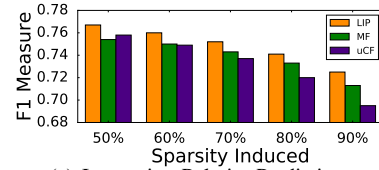


Fig. 3. Results with varied Sparsity Settings.

- *uCF*: User-based collaborative filtering approach where we used the five most similar users (in terms of cosine similarity) based on their helpfulness rating history for making the predictions. For details on collaborative filtering please see [24].
- *MF*: Our low-rank matrix factorization method as shown in Eq. (2).

For link prediction, the representative baselines are presented below and details of the methods can be found in their respective cited work.

- *SSA*: A spectral based method using the signed Laplacian matrix [8] and regularized Laplacian kernel [25] is used. Due to the fact this method was presented for undirected networks, we convert the directed link information by making \mathbf{T} symmetric, thus resulting in an undirected network.
- *HOC-3*: The supervised approach presented in [5] that is based on extracting 23 features; 16 directed triad configurations and 7 node related features.
- *MF*: Low-rank matrix factorization method as shown in Eq. (1), which was first introduced in [4].

In the first experiment, we are able to simulate a ranging sparsity across each user, since we have already limited our attention to a subset of the data that is denser than the original dataset. We remove $x\%$ of the links and interactions for each user and vary x in $\{50, 60, 70, 80, 90\}$.

1) *Experimental Results*: The interaction polarity prediction results can be found in Figure 3(a). Most of the time, we see that the baseline MF method outperforms the user-based collaborative filtering method. Similarly, we have LIP finding significant gains over MF across the levels of sparsity induced. Another thing to mention is that since we had first increased the density of the user-review matrix, it is not until the 80% sparsity that the density of the network drops below that of the original matrix \mathbf{H} .

We report the results of the sparsity experiments for the link prediction in Figure 3(b). LIP and MF obtain much better performance than SSA and HOC-3. We are able to observe that LIP performs comparable to the MF method for the lower sparsity settings, but upon reaching the higher sparsity level, LIP achieves better performance than MF.

From the results in the sparsity experiment, we have seen

LIP’s ability to help alleviate the sparsity problem found in the interaction polarity and link prediction tasks; thus providing evidence that our joint framework is able to partially alleviate the sparsity problem inherent in signed networks. More specifically, we see a significant improvement in the interaction polarity predictions, and increasing improvement for the link prediction with the increase of the sparsity.

C. Cold-Start Experiments

Note that one of the main contributions of this work is the ability of the framework to handle not just the data sparsity problem, but also to help alleviate issues that are commonly faced with cold-start users in signed networks, which are quite common characteristics in these datasets. Therefore to answer the second question, we compare LIP with existing algorithms that are able to handle cold-start users in both of the two prediction tasks.

For this experiment we want to empirically evaluate the robustness of LIP when faced with networks having cold-start users. Note that this is a very difficult problem to overcome due to the fact if there is no knowledge about a user in a certain domain, then it becomes difficult, if not impossible, to make reasonable predictions involving them. However, since LIP is jointly predicting the signed links and user interaction polarities, the opinions formulated in one task can power those in the other task and simultaneously they should be able to gain information for users that previously had none in one of the tasks.

Under the cold-start setting, we choose the following user interaction polarity prediction baselines:

- *RG*: Random guessing method that predicts polarities based on the training data class distribution.
- *AvgG*: The average guessing method (*AvgG*), first calculates the average interaction value found in the entire training set, and then predicts that value for all missing values.
- *MFwRG*: Typical matrix factorization method, but for cold-start users we perform random guessing.

We note that the typical matrix factorization method would not be applicable in this experiment, since if we have no interaction information for a given user, then the latent vectors of such users would never be updated. This would leave the predicted value to be assigned based on the inner product of two randomly initialized vectors. Thus, we modified *MF* by adding the condition that if either of the two users’ vectors have not been updated (i.e., they had no training interaction data and are therefore a cold-start user), then instead of using the inner product as we normally would with *MF* for predicting links, we instead use the *RG* method for that given link.

We compare the proposed framework LIP with the following link prediction baselines:

- *RG*: Randomly guess missing links to be positive or negative based on training data class distribution.
- *MFwRG*: This method has the identical extension for the cold-start users as described in *MFwRG* for the interaction polarity prediction task.

TABLE II

Method	Induced Percent Cold Start Users				
	5%	10%	15%	20%	25%
RG	0.655	0.655	0.655	0.655	0.655
AvgG	0.667	0.667	0.667	0.667	0.667
MFwRG	0.769	0.764	0.754	0.746	0.739
LIP	0.773	0.771	0.769	0.766	0.763

TABLE III

Method	Induced Percent Cold Start Users				
	5%	10%	15%	20%	25%
RG	0.641	0.641	0.641	0.641	0.640
MFwRG	0.848	0.837	0.825	0.813	0.797
LIP	0.860	0.858	0.853	0.848	0.839

For these experiments, we vary the percentage of users that become cold start users in a given task, but do not modify the testing set. We randomly select $x\%$ of the users and remove all their links, then randomly select $x\%$ of the users (who we have not already selected) and remove their interaction information while varying x in $\{5, 10, 15, 20, 25\}$.

1) *Experimental Results*:: Table II holds the results of the cold-start experiments for the interaction polarity prediction task when varying the number of cold start users. The very naive baseline *RG* is just shown to provide a reference for the F1 measure, but the *MFwRG* is expected to perform quite well. In this table, we are able to observe LIP’s superiority over the baseline methods when observing cold-start users. We also see that LIP’s performance as compared to the baselines drastically increases as the number of cold-start users increases, which is extremely intuitive based upon the use of the correlation terms. This is because even if a user has no current helpfulness rating information, LIP is able to transfer information (i.e., their opinions) through the linear mapping matrices \mathbf{W}_O and \mathbf{W}_I and use information that the user had from their link information.

In Table III, we present the link prediction results when varying the amount of cold-start users in the training set. Upon seeing these results the advantages of LIP over the other baseline methods become even more obvious. We note that whenever *MFwRG* has the ability to learn a low dimensional representation for a user, it can then perform the prediction using its learned low dimensional latent vectors. But when there is no link information for a given user, then the user must resort to randomly guessing. Similarly to the interaction polarity prediction task, as the percentage of cold-start users increases, the performance gap in terms of F1 becomes larger in favor of LIP having the best prediction.

D. Discussions

This leads us back to our second question, where we set out to determine if the linking terms based upon the correlated user opinions in LIP are able to provide a transfer of information between the two tasks that ultimately have a user’s opinions in one task power the other. Based upon the results presented in this section, for both the sparsity and the cold-start experiments, we have shown that indeed LIP is able to utilize the inherent correlations behind the opinions expressed in the two tasks to boost the performance in both the prediction tasks simultaneously. Next we present

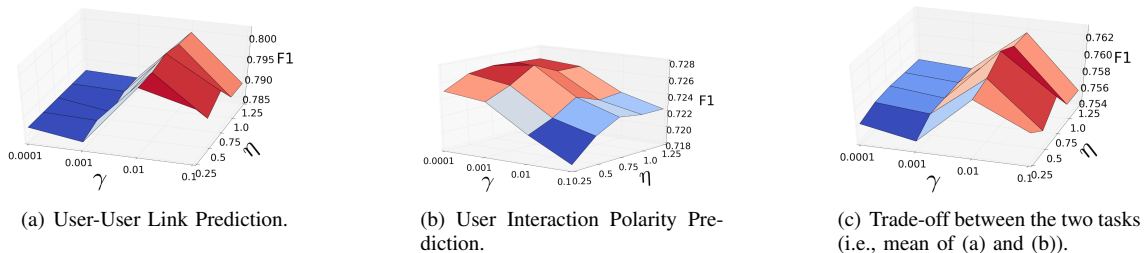


Fig. 4. Performance Variations of LIP on the 90% data sparsity experiment w.r.t. η and γ .

our analysis on the parameters of LIP. We seek to not only to gain a better understanding of the relation between these two prediction tasks (i.e., η), but perhaps even more important in this study, is the focus on γ , since it controlled the amount of opinion information to be transferred from one prediction task to the other.

E. Parameter Analysis for LIP

The parameters η and γ control the balance between optimizing the link prediction and user interaction polarity tasks, and how strongly to keep the two tasks low dimensional representations correlated, respectively. In this subsection, we perform an analysis on how changing these two parameters affects the performance of LIP. We first fix all other parameters (i.e., the regularization parameters β_1, β_2 , and β_3 and dimension sizes K_L and K_I) based upon the best parameters found against our validation set when performing a grid search over the parameter space. We evaluate the performance on all paired (η, γ) values while we vary the value of η as 0.25, 0.5, 0.75, 1.0, 1.25 and γ as 0.0001, 0.001, 0.01, 0.1, providing us with 20 possible combinations. Although the best parameter settings varied between the two above mentioned experiments, we only display one representative from the sparsity user experiment, since we have similar observations in every other experimental setting. We present the analysis on the 90% sparsity experiment since it had the most variation in performance across the different settings.

In Figure 4, we have shown the 3D surfaces for the mentioned combination of parameters. In Figure 4(a), we can see that $\gamma = 0.01$ is shown to clearly be a good region for this parameter, as both to the left and right the performance in terms of F1 drops for the link prediction. However, there is little to no significant difference between the link predictions when varying η in the range provided. It can also be noticed that for the interaction polarity prediction task (seen in Figure 4(b)) the larger η leads to much better performance, which intuitively makes sense because a larger η relates to increasing the weight of how much we were to optimize the interaction polarity prediction as compared to the link prediction task. Unlike what we observed in the link prediction task, the interaction polarity prediction performs better with a smaller γ ; meaning the two tasks have a different preferred weight to be associated with the correlation between the user latent vectors.

Finally Figure 4(c) shows that there is a drastic trade-off between the two tasks. Where if one of the tasks has a large increase in F1, then the other task becomes slightly worse.

Thus to obtain better performance in both tasks, we would want to choose a parameter setting such that the trade-off between the two tasks is balanced. Based on our analysis such a point would have $\gamma = 0.01$, but as for the value of η , there is not a decisive value to choose. Thus, we have shown that the balance between optimizing the two tasks is not very sensitive, although from the figure it appears choosing $\eta = 0.75$ has a slight advantage in both of the two tasks.

VI. RELATED WORK

Although there has been a large number of recent works focused on signed link prediction and even interaction polarity prediction, most of their major drawbacks have been that they optimized each task one at a time.

Previous work on link prediction in signed networks can be split into two primary categories; supervised and unsupervised methods. It was in [5] that the supervised method, HOC-3, was first introduced. They had used the social balance theory to derive 16 features based upon the possible triad configurations and also included 7 additional node features. Later in [17] HOC-3 was extended to higher order cycles, and although it obtained slight improvements, it came at great time complexity costs when the network size becomes large. Thus, it is not as practical for current large real-world networks. The first low-rank approximation method for signed networks was presented in [4], where multiple methods for matrix completion and matrix factorization were discussed.

The literature on the interaction polarity prediction is quite limited in comparison to the classical link prediction task. It was in [16] that the authors had the objective of specifically attempting to predict the rating a user would give the content generated by another user. Unlike our work, they included information about the content of the reviews whereas we have only focused on predictions based upon the network information, although (as mentioned before) we have left this as a future work to include the content information as a means to gain even better prediction results. In [21] they used the interactions for increased performance in recommendations to the users. This achieved better performance over the classical recommender system approaches primarily because they had included the role of users rating reviews as compared to only focusing on the information present in the reviews made by the user themselves. In [20] the authors focused on personalized predictions for review helpfulness were they presented a tensor factorization model. Note that we did not compare with tensor based factorization methods due to the fact they require a higher time and

space complexity and instead we had chosen to use matrix factorization as the base method to extend showing that the correlation between the two tasks are able to provide performance improvements.

VII. CONCLUSION AND FUTURE WORK

In signed networks, users can express their opinions via two activities, i.e., creating signed links and expressing opinions on the content from others. Intuitively, the opinions and behaviors that the users have when performing these two activities online should be related. We first performed an analysis to validate the correlations between the signed links and user interaction polarities from both global and local perspective. Our results show that indeed there is a strong relation between the way users behave in expressing their opinions when performing these two aforementioned activities. We next proposed a joint optimization framework, LIP, for the prediction of signed links and interaction polarities that was built upon having the opinions in one task power the other. This novel framework was able to boost the performance in both prediction tasks when jointly solving the two problems as compared to separately solving them individually. The significance becomes even more important in settings where the social network data is sparse or involves cold-start users. This is due to the fact that LIP is able to partially avoid and mitigate these problems since it can transfer information about users opinions from one problem to another by capturing the correlations between them. Our experiments on a real-world signed network have demonstrated both the effectiveness of LIP and also its robustness to the data sparsity and cold-start problems.

Future work in this domain will be to seek other problems that users might have correlated opinions or behaviors that can be harnessed to increase the performance in multiple tasks simultaneously. We also would like to investigate the underlying dynamics in signed networks that are causing these correlations, or other phenomenon, such as high reciprocity in some networks and not in others. More specifically how reciprocity relates to ways in which users express their opinions and perhaps sometimes even seek revenge.

REFERENCES

- [1] F. Heider, "Attitudes and cognitive organization," *The Journal of psychology*, vol. 21, no. 1, pp. 107–112, 1946.
- [2] D. Cartwright and F. Harary, "Structural balance: a generalization of heider's theory," *Psychological review*, vol. 63, no. 5, p. 277, 1956.
- [3] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," in *Proceedings of the 13th international conference on World Wide Web*. ACM, 2004, pp. 403–412.
- [4] C.-J. Hsieh, K.-Y. Chiang, and I. S. Dhillon, "Low rank modeling of signed networks," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 507–515.
- [5] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 641–650.
- [6] J. Tang, C. Aggarwal, and H. Liu, "Node classification in signed social networks," in *Proceedings of the 2016 SIAM International Conference on Data Mining*. SIAM, 2016, pp. 54–62.
- [7] K.-Y. Chiang, C.-J. Hsieh, N. Natarajan, I. S. Dhillon, and A. Tewari, "Prediction and clustering in signed networks: a local to global perspective," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1177–1213, 2014.
- [8] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, "Spectral analysis of signed graphs for clustering, prediction and visualization," in *Proceedings of the 2010 SIAM International Conference on Data Mining*. SIAM, 2010, pp. 559–570.
- [9] T. Sharma, "Finding communities in weighted signed social networks," in *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*. IEEE, 2012, pp. 978–982.
- [10] P. Victor, C. Cornelis, M. De Cock, and A. Teredesai, "Trust-and distrust-based recommendations for controversial reviews," in *Web Science Conference (WebSci'09: Society On-Line)*, no. 161, 2009.
- [11] H. Ma, M. R. Lyu, and I. King, "Learning to recommend with trust and distrust relationships," in *Proceedings of the third ACM conference on Recommender systems*. ACM, 2009, pp. 189–196.
- [12] R. Forsati, M. Mahdavi, M. Shamsfard, and M. Sarwat, "Matrix factorization with explicit trust and distrust side information for improved social recommendation," *ACM Transactions on Information Systems (TOIS)*, vol. 32, no. 4, p. 17, 2014.
- [13] J. Tang, S. Chang, C. Aggarwal, and H. Liu, "Negative link prediction in social media," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. ACM, 2015, pp. 87–96.
- [14] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, "A Survey of Signed Network Mining in Social Media," *ACM Comput. Surv. Article*, vol. 9, no. 39, 2016.
- [15] J. Tang, X. Hu, and H. Liu, "Is distrust the negation of trust?: the value of distrust in social media," in *Proceedings of the 25th ACM conference on Hypertext and social media*. ACM, 2014, pp. 148–157.
- [16] J. Tang, H. Gao, X. Hu, and H. Liu, "Context-aware review helpfulness rating prediction," in *Proceedings of the 7th ACM conference on Recommender systems*. ACM, 2013, pp. 1–8.
- [17] K.-Y. Chiang, N. Natarajan, A. Tewari, and I. S. Dhillon, "Exploiting longer cycles for link prediction in signed networks," in *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 1157–1162.
- [18] P. Symeonidis and N. Mantas, "Spectral clustering for link prediction in social networks with positive and negative links," *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1433–1447, 2013.
- [19] S. Moghaddam, M. Jamali, and M. Ester, "Review recommendation: personalized prediction of the quality of online reviews," in *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 2249–2252.
- [20] J. M. Moghaddam, Samaneh and M. Ester, "Etf: extended tensor factorization model for personalizing prediction of review helpfulness," in *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, 2012, pp. 163–172.
- [21] S. Wang, J. Tang, and H. Liu, "Toward dual roles of users in recommender systems," in *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*. ACM, 2015, pp. 1651–1660.
- [22] P. Agrawal, V. K. Garg, and R. Narayanam, "Link label prediction in signed social networks," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, ser. IJCAI '13. AAAI Press, 2013, pp. 2591–2597. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2540128.2540502>
- [23] R. Forsati, I. Barjasteh, F. Masrour, A.-H. Esfahanian, and H. Radha, "Pushtrust: An efficient recommendation algorithm by leveraging trust and distrust relations," in *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 2015, pp. 51–58.
- [24] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in artificial intelligence*, vol. 2009, p. 4, 2009.
- [25] T. Ito, M. Shimbo, T. Kudo, and Y. Matsumoto, "Application of kernels to link analysis," in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005, pp. 586–592.
- [26] P. Symeonidis and E. Tiakas, "Transitive node similarity: predicting and recommending links in signed social networks," *World Wide Web*, vol. 17, no. 4, pp. 743–776, 2014.