Predictability of Distrust with Interaction Data

Jiliang Tang[‡], Xia Hu[‡], Yi Chang[†] and Huan Liu[‡]
[‡]Computer Science and Engineering,Arizona State University,Tempe, AZ, USA
{Jiliang.Tang, Xia.Hu, Huan.Liu}@asu.edu
[†]Yahoo!Labs, Yahoo!Inc, Sunnyvale,CA, USA
yichang@yahoo-inc.com

ABSTRACT

Trust plays a crucial role in helping users collect reliable information in an online world, and has attracted more and more attention in research communities lately. As a conceptual counterpart of trust, distrust can be as important as trust. However, distrust is rarely studied in social media because distrust information is usually unavailable. The value of distrust has been widely recognized in social sciences and recent work shows that distrust can benefit various online applications in social media. In this work, we investigate whether we can obtain distrust information via learning when it is not directly available, and propose to study a novel problem - predicting distrust using pervasively available interaction data in an online world. In particular, we analyze interaction data, provide a principled way to mathematically incorporate interaction data in a novel framework dTrust to predict distrust information. Experimental results using real-world data show that distrust information is predictable with interaction data by the proposed framework dTrust. Further experiments are conducted to gain a deep understand on which factors contribute to the effectiveness of the proposed framework.

Categories and Subject Descriptors

H3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Information filtering*

General Terms

Algorithms; Design; Experimentation

Keywords

Distrust in Social Media; Predictability of Distrust; Interaction Data; Balance Theory

1. INTRODUCTION

Trust plays an important role in helping online user obtain reliable information and has attracted increasing attention in recent years [9, 36, 37]. The availability of trust information helps improve the performance of various applications in social media such as recommendation [9, 22], finding high-quality user generated content [21], and viral marketing [31]. However, as a conceptual counterpart of trust, distrust is rarely studied in the online world because distrust information is usually unavailable in the online world.

It is suggested in research [16, 11] that trust is a desired property while distrust is an unwanted one for an online social community. Therefore, various online services such as Ciao[1], eBay[2] and Epinions[3] implement trust mechanisms to help users to better use their services, but few of them allow online users to specify distrust relations. As we learn from social sciences [26, 16], distrust can be as important as trust. Both trust and distrust can help a decision maker reduce the uncertainty and vulnerability associated with decision consequences [4], and sometimes distrust can play a critical role in consumer decisions [34, 26]. Many social scientists believe that distrust is not simply the negation of trust [18, 15]. This can be quickly verified using transitivity[4] and balance theory [2, 12]. As we know, transitivity is an important property of trust, i.e., if u_i trusts u_j and u_j trusts u_k, it is likely for u_i to trust u_k. If distrust was the negation of trust (or distrust is equivalent to low trust) [33], according to transitivity, it would be true that if u_i distrusts u_j and u_j distrusts u_k, it is likely that u_i distrusts u_k. But the transitivity of distrust would violate balance theory[5].

The lack of distrust research could lead to a biased estimate of the effect of trust, and distrust information has added value on trust [38]. The availability of distrust information can benefit various applications in social media. For example, a small amount of distrust information can make remarkable improvement in trust prediction [38], and in e-commerce, users might or might not accept recommendations from their trusted users, but will certainly exclude recommendations from their distrusted users [26, 41]. Genuine distrust information tends to be more noticeable and credible, and weighed more than trust information of a similar magnitude, therefore, distrust information can be critical in social media applications. Given the fact that distrust information is usually unavailable but important, in this work, we investigate if distrust information can be obtained via learning from social media data.

CIKM'14, November 3–7, 2014, Shanghai, China.
Copyright 2014 ACM 978-1-4503-2598-1/14/11 ...$15.00.
http://dx.doi.org/10.1145/2661829.2661988.

[1]http://www.ciao.co.uk/
[2]http://www.ebay.com/
[3]http://www.epinions.com/
[4]A relation R is transitive if u_iRu_j and u_jRu_k, then u_iRu_k.
[5]If u_i *distrusts* u_j and u_j *distrusts* u_k, then u_i *trusts* u_k.

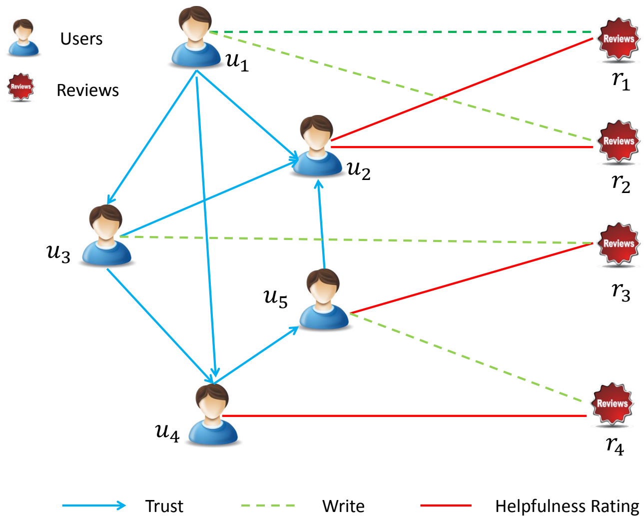


Figure 1: An Illustration of Interaction Data in Product Review Sites.

In the physical world, people are likely to share trust with others while hide their distrust. However, we still can tell distrust by observing their interactions. With this intuition, interaction data in social media might be helpful to obtain distrust information when it is not directly available. Interaction data is pervasively available in social media. Figure 1 illustrates interaction data available in product review sites like Epinions. Trust relations among users (e.g., $\{u_1 \rightarrow u_2\}$) are available. Users can give helpfulness ratings to reviews written by other users. For example, u_2 rates the helpfulness of reviews r_1 and r_2 written by u_1. Users can rate a review with various scores to indicate its helpfulness from "not helpful" to "very helpful". In this paper, we propose to turn the search of distrust information into a problem of predictability of distrust using interaction data. In essence, we investigate - (1) how to exploit interaction data? and (2) how to model interaction data mathematically to predict distrust relations? Our solutions to these two challenges result in a novel framework dTrust for the distrust prediction problem. Our main contributions are summarized below,

- Provide a principled way to enable the predictability of distrust by exploiting interaction data;
- Propose a novel framework dTrust to predict distrust relations by incorporating interaction data mathematically; and
- Evaluate the proposed framework dTrust in real-world data to understand the working of dTrust.

The rest of paper is organized as follows. The studied problem is formally defined in Section 2. In Section 3, we describe the datasets and the importance of interactions in the predictability of distrust relations. In Section 4, we introduce the way to incorporate interaction data and the proposed framework dTrust to predict distrust relations with interaction data. Section 5 presents experimental results with discussions. Section 6 reviews related work. Finally, Section 7 concludes this study with future work.

2. PROBLEM STATEMENT

In this paper, we study the problem in the context of product review sites, however, the proposed framework is general and can be applied to other sites implementing trust mechanisms. Let $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ and $\mathcal{R} = \{r_1, r_2, \ldots, r_m\}$ be the sets of users and reviews respectively, where n is the number of users and m is the number of reviews. We use $\mathbf{T} \in \mathbb{R}^{n\times n}$, $\mathbf{D} \in \mathbb{R}^{n\times n}$, $\mathbf{P} \in \mathbb{R}^{n\times m}$, and $\mathbf{R} \in \mathbb{R}^{n\times m}$ to denote user-user trust relations, user-user distrust relations, user-review authorship relations, and user-review helpfulness ratings at time t, respectively. $\mathbf{T}_{ij} = 1$ (or $\mathbf{D}_{ij} = 1$) if u_i trusts (or distrusts) u_j, zero otherwise. $\mathbf{P}_{ij} = 1$ if u_i writes r_j, zero otherwise. If u_i rates the helpfulness of r_j, $\mathbf{R}_{ij}$ is the helpfulness rating score, and we use the symbol "?" to denote $\mathbf{R}_{ij}$ if u_i does not rate r_j.

The availability of trust information $\mathbf{T}$ encourages many trust-related applications, and trust prediction is one of the most important and popular applications. Assume that $\mathbf{T}_t \in \mathbb{R}^{n\times n}$ denotes user-user trust relations established after the time t, trust prediction aims to develop a predictor f to predict new relations $\mathbf{T}_t$ using old relations $\mathbf{T}$,

$$f : \{\mathbf{T}\} \rightarrow \{\mathbf{T}_t\}. \tag{1}$$

When both distrust $\mathbf{D}$ and trust $\mathbf{T}$ information are available, trust and distrust prediction problem is extensively studied to predict new trust and distrust relation as

$$f : \{\mathbf{T}, \mathbf{D}\} \rightarrow \{\mathbf{T}_t, \mathbf{D}_t\}, \tag{2}$$

where $\mathbf{D}_t \in \mathbb{R}^{n\times n}$ denotes user-user distrust relations established after the time t. Trust and distrust prediction problem is to develop a predictor f to predict new trust and distrust relations $\{\mathbf{T}_n, \mathbf{D}_n\}$ using old trust and distrust relations $\{\mathbf{T}, \mathbf{D}\}$.

Given the unavailability and importance of distrust information, we ask whether distrust information can be learned from interaction data. We formally define it as: *given user-user trust relations* $\mathbf{T}$*, user-review authorship relations* $\mathbf{P}$ *and user-review helpfulness ratings* $\mathbf{R}$*, we aim to develop a predictor* f *to predict distrust relations* $\mathbf{D}$ *with* $\mathbf{T}$*,* $\mathbf{P}$ *and* $\mathbf{R}$,

$$f : \{\mathbf{T}, \mathbf{P}, \mathbf{R}\} \rightarrow \{\mathbf{D}\}. \tag{3}$$

3. DATA ANALYSIS

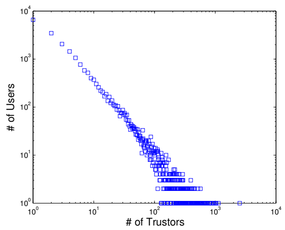
In this section, we first introduce the dataset we used for this study, and then provide our solution to the first challenge - how to exploit interaction data.

3.1 Dataset

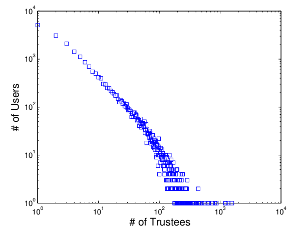
Trust mechanisms are implemented by various online services; however, few of them allow uses to establish distrust relations. Although the product review site Epinions allows users to trust and distrust other users, distrust relations are unavailable to the public. For the research purpose, a dataset with distrust relations was given by Epinions staff [24]. We preprocess the data by filtering users without any trust and distrust relations. This dataset includes trust and distrust relations, user-review authorship relations and user-review helpfulness ratings. Note that the availability of distrust relations in this dataset serves as the ground truth for only analysis and evaluation purpose, which are not used in the learning process of the proposed framework. The statistics of the dataset are shown in Table 1.

Table 1: Statistics of the Epinions Dataset.

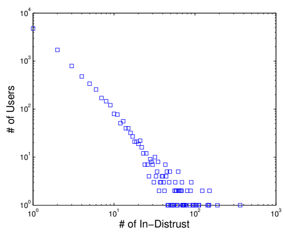
# of Users	30,455
# of Trust Relations	363,773
# of Distrust Relations	46,196
# of Reviews	1,197,816
# of Helpfulness Ratings	10,341,893
Avg of Helpfulness Rating Score	4.7129



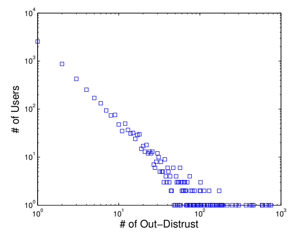
(a) Indegree-Trust



(b) Outdegree-Trust



(c) Indegree-Distrust



(d) Outdegree-Distrust

Figure 2: The Distributions of Indegree and Outdegree of Trust and Distrust Relations.

We compute the number of trust and distrust relations each user receives and creates, and these distributions are shown in Figure 2. The distributions for both trust and distrust suggest a power-law-like distribution that is typical in social networks. Users in Epinions can specify a score from 1 to 6 to indicate the helpfulness of a review from "not helpful" to "very helpful". We investigate the helpfulness rating distributions and find that more than 70% of users give a score 4 or 5 with an average score of 4.7129. In the following subsection, we investigate the correlation between interactions and distrust relations.

3.2 Analysis on Interaction Data

Users can participate in various online activities such as liking, commenting or rating, which produces rich interaction data in social media. Users can perform negative interactions to other users by disliking, giving negative comments or negative ratings on their generated content. In the context of product review sites, a user can rate reviews written by another user not helpful, which shows disagreement and antagonism toward the user. It is reasonable to surmise that negative interactions (e.g., "not helpful" ratings) might be correlated to their distrust relations. In this subsection, we study the correlations between negative interactions and distrust relations to seek a solution to the first challenge.

In Epinions, users can rate reviews with scores 1 to 6 to indicate their helpfulness. In this study, we consider scores less than 3 as not helpful ratings (or negative interactions).

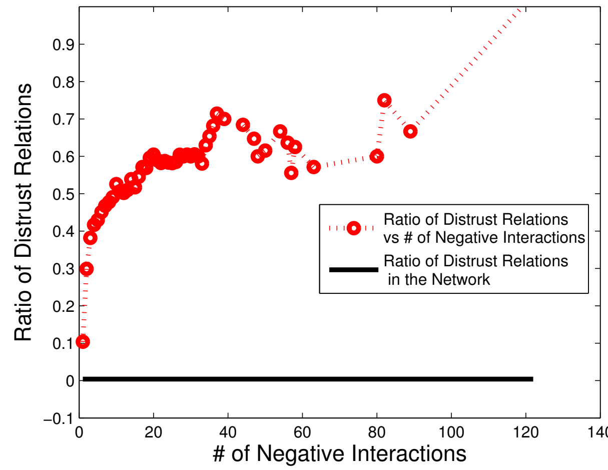


Figure 3: # Negative Interactions vs. Distrust Relations.

We calculate a matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ from user-review authorship relations $\mathbf{P}$ and user-review helpfulness ratings $\mathbf{R}$ where $\mathbf{Q}_{ij}$ is the number of negative interactions from u_i to u_j.

To study the correlation between negative interactions and distrust relations, we try to answer the question - *are two users with negative interactions more likely to have a distrust relation than two randomly chosen users?*. Let $\mathcal{H} = \{\langle u_i, u_j\rangle | \mathbf{Q}_{ij} > 0\}$ be the set of pairs of users with negative interactions. For each pair of users $\langle u_i, u_j\rangle$, we use a_n and b_r to indicate whether there exist distrust relations for $\langle u_i, u_j\rangle$ and $\langle u_i, u_k\rangle$ respectively, where u_k is a randomly chosen user. If u_i distrusts u_j, $a_n = 1$ and zero otherwise; while if u_i distrusts u_k, $b_r = 1$ and zero otherwise. For all pairs in $\mathcal{H}$, we obtain two vectors, $\mathbf{a}$ and $\mathbf{b}$. $\mathbf{a}$ is the set of a_n, while $\mathbf{b}$ is the set of b_r. We conduct a two-sample t-test on $\mathbf{a}$ and $\mathbf{b}$. The null hypothesis and the alternative hypothesis are defined as

$$H_0 : \mathbf{a} \leq \mathbf{b}, \quad H_1 : \mathbf{a} > \mathbf{b}. \tag{4}$$

The null hypothesis is rejected at significance level $\alpha = 0.01$ with p-value of 7.12e-147. Evidence from t-test suggests a positive answer to the question: *there is a strong correlation between negative interactions and distrust relations, and users with negative interactions are likely to have distrust relations.*

We define $\mathcal{K}$ as the set of unique non-zero numbers of negative interactions in $\mathbf{Q}$ as

$$\mathcal{K} = \{K | K \neq 0 \wedge \exists_{\langle u_i, u_j\rangle} \mathbf{Q}_{ij} = K\} \tag{5}$$

To study the impact of the number of negative interactions on the correlation, we use $\mathcal{S}_K$ to denote the set of pairs of users $\langle u_i, u_j\rangle$ where the number of negative interactions from u_i to u_j is no less than $K \in \mathcal{K}$, which is formally defined as

$$\mathcal{S}_K = \{\langle u_i, u_j\rangle | \mathbf{Q}_{ij} \geq K\}. \tag{6}$$

We further define that $\mathcal{D}_K$ is the set of pairs with distrust relations in $\mathcal{S}_K$ as

$$\mathcal{D}_K = \{\langle u_i, u_j\rangle | (\mathbf{Q}_{ij} \geq K) \wedge \mathbf{D}_{ij} = 1\}. \tag{7}$$

Then we calculate p_K as the ratio of pairs with distrust relations $\mathcal{D}_K$ in $\mathcal{S}_K$ as

$$p_K = \frac{|\mathcal{D}_K|}{|\mathcal{S}_K|}. \tag{8}$$

The distribution of ratios of distrust relations p_Ks with respect to $\mathcal{K}$ is shown in Figure 3. Note that the black solid line in the figure denotes the ratio of distrust relations r_{dis} in the distrust network. For all K, p_Ks are much larger than r_{dis}, which further suggests the correlation between negative interactions and distrust relations. With the increase of K, the ratios p_Ks tend to increase, which indicates that the more negative interactions two users have, the more likely a distrust relation exists between them.

4. OUR FRAMEWORK - dTrust

In the last section, we find a strong correlation between negative interactions and distrust relations, and reveal the impact of the number of negative interactions on the correlation. In this section, we first introduce a way to model interaction data by capturing the correlation, and then present the proposed framework dTrust with its optimization algorithm, which provides the solution to the second challenge.

4.1 Modeling Interaction Data

To model interaction data, we try to capture the correlation with findings from the previous section. We first divide all n^2 pairs of users into three groups $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3\}$, and their definitions are stated as:

- $\mathcal{G}_1$ contains pairs of users with trust relations as
$$\mathcal{G}_1 = \{\langle u_i, u_j\rangle | \mathbf{T}_{ij} = 1\}, \tag{9}$$
- $\mathcal{G}_2$ is the set of pairs of users without trust relations but with negative interactions as
$$\mathcal{G}_2 = \{\langle u_i, u_j\rangle | \mathbf{T}_{ij} = 0 \wedge \mathbf{Q}_{ij} > 0\}, \tag{10}$$
- $\mathcal{G}_3$ includes the remaining pairs as,
$$\mathcal{G}_3 = \{\langle u_i, u_j\rangle | \langle u_i, u_j\rangle \notin (\mathcal{G}_1 \cup \mathcal{G}_2)\}. \tag{11}$$

From above definitions, we can see that $\mathcal{G}_1$ and $\mathcal{G}_2$ correspond to the set of pairs of users with trust relations and negative interactions, respectively. Based on these three groups, we introduce a matrix $\mathbf{F} \in \mathbb{R}^{n\times n}$ to represent user-user trust relations and pseudo distrust relations from interaction data, and the entities of $\mathbf{F}$ are defined as follows:

- For $\langle u_i, u_j\rangle \in \mathcal{G}_1$, we set $\mathbf{F}_{ij} = 1$ since u_i trusts u_j;
- For $\langle u_i, u_j\rangle \in \mathcal{G}_2$, u_i gives negative interactions to u_j, and according to the correlation between negative interactions and trust relations, u_i is likely to distrust u_j; hence, we assign a pseudo distrust relation from u_i to u_j by setting $\mathbf{F}_{ij} = -1$;
- We do not have evidence of possible relations for $\langle u_i, u_j\rangle \in \mathcal{G}_3$, therefore we set $\mathbf{F}_{ij} = 0$ as a missing relation.

The entities of $\mathbf{F}$ are formally defined as follows:

$$\mathbf{F}_{ij} = \begin{cases} 1 & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_1 \\ -1 & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_2 \\ 0 & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_3 \end{cases}. \tag{12}$$

The values in the user-user trust and pseudo distrust matrix $\mathbf{F}$ may be not equally reliable. For example, $\mathbf{F}_{ij}$ for $\langle u_i, u_j\rangle \in \mathcal{G}_1$ is very reliable since we observe trust relations, while values of pairs in $\mathcal{G}_2$ with more negative interactions are more reliable based on our previous finding of the impact of the number of negative interactions on the correlations - the more negative interactions two users have, the more likely a distrust relation exists between them. Therefore, we define a weight matrix $\mathbf{W} \in \mathbb{R}^{n\times n}$ where $\mathbf{W}_{ij} \in [0, 1]$ is a weight to indicate the reliability of $\mathbf{F}_{ij}$. Next we define the weight matrix as

- We observe trust relations for pairs in $\mathcal{G}_1$; hence for $\langle u_i, u_j\rangle \in \mathcal{G}_1$, we set $\mathbf{W}_{ij} = 1$;
- Our previous finding reveals that the more negative interactions two users have, the more likely a distrust relation between them exists; hence for $\langle u_i, u_j\rangle \in \mathcal{G}_2$, $\mathbf{W}_{ij}$ is defined as a function of the number of negative interactions as $\mathbf{W}_{ij} = g(\mathbf{G}_{ij})$. The function $g(x)$ has following properties - (1) x is a positive integer; (2) $g(x) \in [0, 1]$; and (3) $g(x)$ is non-decreasing function of x; and
- We set $\mathbf{W}_{ij}$ to be a constant $c \in [0, 1]$ for $\langle u_i, u_j\rangle \in \mathcal{G}_3$.

We empirically find the following definition of $g(x)$ works well in this work:

$$g(x) = 1 - \frac{1}{log(x + 1)}. \tag{13}$$

where $g(x)$ is a non-decreasing function of x. In our problem, x is a positive integer therefore $x + 1$ can guarantee that $(1 - \frac{1}{log(x+1)}) \in [0, 1]$.

The weight matrix $\mathbf{W}$ is formally defined as,

$$\mathbf{W}_{ij} = \begin{cases} 1 & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_1 \\ g(\mathbf{Q}_{ij}) & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_2 \\ c & \text{if } \langle u_i, u_j\rangle \in \mathcal{G}_3 \end{cases}. \tag{14}$$

With the user-user trust and pseudo distrust relations $\mathbf{F}$ and its weight matrix $\mathbf{W}$, our problem can boil down to a special trust and distrust prediction problem with trust and pseudo distrust relations. Therefore we can choose a representative trust and distrust prediction algorithm as our basic algorithm. In [35], a matrix factorization framework is proposed to predict trust relations based on $\mathbf{T}$ as

$$\min_{\mathbf{U},\mathbf{H}} \quad \|\mathbf{T} - \mathbf{U}\mathbf{H}\mathbf{U}^\top\|_F^2 + \alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{H}\|_F^2), \tag{15}$$

where $\mathbf{U} \in \mathbb{R}^{n\times d}$ with $d \ll n$ is the user preference matrix and $\mathbf{T}_{ij}$ is modeled as the correlation between the preferences of u_i and u_j by $\mathbf{H}$ as $\mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top$. The term $\alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{H}\|_F^2)$ is added to avoid over-fitting. The framework can be directly extended for trust and distrust prediction by representing a distrust relation as -1 in $\mathbf{T}$ [13]. In this paper, we choose it as the basic algorithm. However, note that we could also choose other algorithms such as trust and distrust propagation [10] as the basic algorithm and we would like to leave this investigation for our future work. We may not directly apply Eq. (15) to our problem since the values in $\mathbf{F}$ may not be reliable. We modify Eq. (15) and the new formulation with the user-user trust and pseudo distrust relations $\mathbf{F}$ and its weight matrix $\mathbf{W}$ is to solve the following optimization problem,

$$\min_{\mathbf{U},\mathbf{H}} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\mathbf{W}_{ij}(\mathbf{F}_{ij} - \mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top)\right)^2 + \alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{H}\|_F^2), \tag{16}$$

where the formation in Eq. (16) allows us to consider the reliability of values in $\mathbf{F}$ by $\mathbf{W}$. The contribution of $\mathbf{F}_{ij}$ to the learning process is controlled by $\mathbf{W}_{ij}$. A large value of $\mathbf{W}_{ij}$, indicating the high reliability of $\mathbf{F}_{ij}$, will force $\mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top$ to tightly fit $\mathbf{F}_{ij}$ while $\mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top$ will loosely approximate $\mathbf{F}_{ij}$ when $\mathbf{W}_{ij}$ is small.

To model interaction data, we introduce the concept of pseudo distrust relations and the significance is three-fold. First, it provides a way to model interaction data. Second, it helps us boil down the studied problem into a special trust and distrust prediction problem. Finally it allows us to exploit some social theories for signed networks since the introduction of pseudo distrust relations converts the trust unsigned network into a signed trust and pseudo distrust network. In the following subsection, we will introduce how to model one of the most important and popular social theories for signed networks balance theory.

4.2 Modeling Balance Theory

We use s_{ij} to denote the sign of the relation between u_i and u_j where $s_{ij} = 1$ (or $s_{ij} = -1$) if we observe a trust relation (or a distrust relation) between u_i and u_j. With these notations, balance theory suggests that a triad $\langle u_i, u_j, u_k\rangle$ is balanced if

- $s_{ij} = 1$ and $s_{jk} = 1$, then $s_{ik} = 1$; or
- $s_{ij} = -1$ and $s_{jk} = -1$, then $s_{ik} = 1$.

Note that balance theory is proposed for undirected networks and following a common practice [17], we ignore the directions of relations, and only consider the signs of relations (i.e., trust and distrust) when we apply balance theory to trust and pseudo distrust relations.

For a triad $\langle u_i, u_j, u_k\rangle$, there are four possible sign combinations (+,+,+) (+,+,-) (-,-,+) and (-,-,-), while only (+,+,+) and (-,-,+) are balanced. We examine all triads in the studied dataset and find that more than 90% of them are balanced, which is consistent with observations in [17]. This result suggests that balance theory is a principle to understand the formation of trust and distrust relations. The introduction of pseudo distrust relations enables us to exploit balance theory, while exploiting balance theory in turn may help us mitigate the effects of unreliability of pseudo distrust relations, and potentially improves the distrust prediction performance.

There are three common ways to exploit social theories in social media mining including feature engineering, constraint generating, and objective defining [39]. In this work, we choose objective defining to model balance theory. For each user u_i, we introduce a one-dimensional latent factor r_i and we further assume that the trust or distrust relation between u_i and u_j due to the effect of balance theory is modeled as [42],

$$\mathbf{F}_{ij} = r_i r_j, \tag{17}$$

Next we will prove that Eq. (17) can capture balance theory with the following theorem:

THEOREM 4.1. *Eq. (17) can capture balance theory. That is to say, with Eq. (17), we can have*

- *Case 1: If sign*$(\mathbf{F}_{ij}) = 1$ *and sign*$(\mathbf{F}_{jk}) = 1$*, we can prove that sign*$(\mathbf{F}_{ik}) = 1$.
- *Case 2: If sign*$(\mathbf{F}_{ij}) = -1$ *and sign*$(\mathbf{F}_{jk}) = -1$*, we can prove that sign*$(\mathbf{F}_{ik}) = 1$.

PROOF. Let us first prove *Case 1.* If $sign(\mathbf{F}_{ij}) = 1$ and $sign(\mathbf{F}_{jk}) = 1$, we have $sign(r_ir_j) = 1$ and $sign(r_jr_k) = 1$; by multiplying $sign(r_ir_j)$ and $sign(r_jr_k)$, we have $sign(r_ir_jr_jr_k) = 1$. Since $sign(r_j^2) = 1$, we get $sign(r_ir_k) = 1$, i.e., $sign(\mathbf{F}_{ik}) = 1$.

We can use a similar process to prove *Case 2.* If $sign(\mathbf{F}_{ij}) = -1$ and $sign(\mathbf{F}_{jk}) = -1$, we have $sign(r_ir_j) = -1$ and $sign(r_jr_k) = -1$; by multiplying $sign(r_ir_j)$ and $sign(r_jr_k)$, we have $sign(r_ir_jr_jr_k) = 1$. Since $sign(r_j^2) = 1$, we get $sign(r_ir_k) = 1$, i.e., $sign(\mathbf{F}_{ik}) = 1$, which completes the proof. □

With the solutions to both challenges in the introduction section, next we will introduce the proposed framework with its optimization algorithm.

4.3 An Optimization Algorithm for dTrust

In Eq. (16), $\mathbf{F}_{ij}$ is modeled as $\mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top$, and it is modeled as r_ir_j due to the effect of balance theory. When we consider both, $\mathbf{F}_{ij}$ can be modeled by combining Eqs. (16) and (17) as

$$\mathbf{F}_{ij} = \mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top + \lambda r_i r_j, \tag{18}$$

where λ is introduced to control the contribution from balance theory. Then the proposed framework dTrust is to solve the following optimization problem,

$$\begin{aligned}\min_{\mathbf{U},\mathbf{H},\mathbf{r}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\mathbf{W}_{ij}(\mathbf{F}_{ij} - \mathbf{U}_i\mathbf{H}\mathbf{U}_j^\top - \lambda r_i r_j)\right)^2 \\ & + \alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{H}\|_F^2 + \|\mathbf{r}\|_2^2),\end{aligned} \tag{19}$$

where $\mathbf{r} = [r_1, r_2, \ldots, r_n]^\top$ and the term $\eta\|\mathbf{r}\|_2^2$ is introduced to avoid over-fitting. Eq. (19) can be rewritten to its matrix form as

$$\begin{aligned}\min_{\mathbf{U},\mathbf{H},\mathbf{r}} \quad & \|\mathbf{W} \odot (\mathbf{F} - \mathbf{U}\mathbf{H}\mathbf{U}^\top - \lambda\mathbf{r}\mathbf{r}^\top)\|_F^2 \\ & + \alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{r}\|_2^2),\end{aligned} \tag{20}$$

where $\odot$ is the Hadamard product where $(\mathbf{X}\odot\mathbf{Y})_{ij} = \mathbf{X}_{ij} \times \mathbf{Y}_{ij}$ for any two matrices $\mathbf{X}$ and $\mathbf{Y}$ with the same size.

Set $\mathbf{A} = \mathbf{F} - \lambda\mathbf{r}\mathbf{r}^\top$ and let $\mathcal{L}$ contain terms related to $\mathbf{U}$ and $\mathbf{H}$ in the objective function $\mathcal{J}$ of Eq. (20), which can be rewritten as,

$$\begin{aligned}\mathcal{L} = Tr(&-2(\mathbf{W}\odot\mathbf{W}\odot\mathbf{A})\mathbf{U}\mathbf{H}^\top\mathbf{U}^\top + (\mathbf{W}\odot\mathbf{W}\odot\mathbf{U}\mathbf{H}\mathbf{U}^\top) \\ & \mathbf{U}\mathbf{H}^\top\mathbf{U}^\top) + \alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)\end{aligned} \tag{21}$$

the partial derivations of $\mathbf{U}$ and $\mathbf{H}$ with respective to $\mathcal{J}$ can be obtained from $\mathcal{L}$ are

$$\begin{aligned}
&\frac{1}{2}\frac{\partial \mathcal{J}}{\partial \mathbf{U}} = \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = \\
&-(\mathbf{W}\odot\mathbf{W}\odot\mathbf{A})\mathbf{U}\mathbf{H}^\top - (\mathbf{W}\odot\mathbf{W}\odot\mathbf{A})^\top\mathbf{U}\mathbf{H} + \alpha\mathbf{U} \\
&+(\mathbf{W}\odot\mathbf{W}\odot\mathbf{U}\mathbf{H}\mathbf{U}^\top)\mathbf{U}\mathbf{H}^\top + (\mathbf{W}\odot\mathbf{W}\odot\mathbf{U}\mathbf{H}\mathbf{U}^\top)^\top\mathbf{U}\mathbf{H}, \\
&\frac{1}{2}\frac{\partial \mathcal{J}}{\partial \mathbf{H}} = \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \mathbf{H}} = \\
&-\mathbf{U}^\top(\mathbf{W}\odot\mathbf{W}\odot\mathbf{A})\mathbf{U} + \mathbf{U}^\top(\mathbf{W}\odot\mathbf{W}\odot\mathbf{U}\mathbf{H}\mathbf{U}^\top)\mathbf{U} + \alpha\mathbf{H}
\end{aligned} \tag{22}$$

Set $\mathbf{B} = \mathbf{G} - \mathbf{U}\mathbf{H}\mathbf{U}^\top$ and let $\mathcal{L}_r$ contain terms related to $\mathbf{r}$ in $\mathcal{J}$, which can be rewritten as,

$$\begin{aligned}
\mathcal{L}_r = Tr(&-2\lambda(\mathbf{W}\odot\mathbf{W}\odot\mathbf{B})\mathbf{r}\mathbf{r}^\top \\
&+\lambda^2(\mathbf{W}\odot\mathbf{W}\odot\mathbf{r}\mathbf{r}^\top)\mathbf{r}\mathbf{r}^\top) + \alpha\|\mathbf{r}\|_2^2
\end{aligned} \tag{23}$$

then the partial derivation of $\mathbf{r}$ with respect to $\mathcal{J}$ is

$$\begin{aligned}
&\frac{1}{2}\frac{\partial \mathcal{J}}{\partial \mathbf{r}} = \frac{1}{2}\frac{\partial \mathcal{L}_r}{\partial \mathbf{r}} \\
&-\lambda(\mathbf{W}\odot\mathbf{W}\odot\mathbf{B})\mathbf{r} - \lambda(\mathbf{W}\odot\mathbf{W}\odot\mathbf{B})^\top\mathbf{r} + \alpha\mathbf{r} \\
&+\lambda^2(\mathbf{W}\odot\mathbf{W}\odot\mathbf{r}\mathbf{r}^\top)\mathbf{r} + \lambda^2(\mathbf{W}\odot\mathbf{W}\odot\mathbf{r}\mathbf{r}^\top)^\top\mathbf{r}
\end{aligned} \tag{24}$$

With the partial derivations of $\mathbf{U}$, $\mathbf{H}$, and $\mathbf{r}$, a optimal solution of the objective function in Eq. (20) can be obtained through a gradient decent optimization method as shown in Algorithm 1.

Algorithm 1: The Proposed Framework dTrust.

Input : User-user trust relations $\mathbf{T}$, user-review authorship relations $\mathbf{P}$, user-review helpfulness ratings $\mathbf{R}$, $\{d, \lambda\}$.
Output : A ranking list of pairs of users.
1: Construct $\mathbf{W}$ and $\mathbf{F}$ from $\mathbf{T}$, $\mathbf{P}$, and $\mathbf{R}$
2: Initialize $\mathbf{U}$, $\mathbf{H}$ and $\mathbf{r}$ randomly
3: **while** Not convergent **do**
4: Calculate $\frac{\partial \mathcal{J}}{\partial \mathbf{U}}$, $\frac{\partial \mathcal{J}}{\partial \mathbf{H}}$ and $\frac{\partial \mathcal{J}}{\partial \mathbf{r}}$
5: Update $\mathbf{U} \leftarrow \mathbf{U} - \gamma_u \frac{\partial \mathcal{J}}{\partial \mathbf{U}}$
6: Update $\mathbf{H} \leftarrow \mathbf{H} - \gamma_h \frac{\partial \mathcal{J}}{\partial \mathbf{H}}$
7: Update $\mathbf{r} \leftarrow \mathbf{r} - \gamma_r \frac{\partial \mathcal{J}}{\partial \mathbf{r}}$
8: **end while**
9: Set $\hat{\mathbf{F}} = \mathbf{U}\mathbf{H}\mathbf{U}^\top + \lambda\mathbf{r}\mathbf{r}^\top$
10: Set $\mathcal{D} = \{\langle u_i, u_j\rangle | sign(\hat{\mathbf{F}}_{ij}) = -1\}$
11: Ranking pairs of users in $\mathcal{D}$ (e.g.,$\langle u_i, u_j\rangle$) according to $|\hat{\mathbf{F}}|$ (e.g., $|\hat{\mathbf{F}}_{ij}|$) in a descending order

Next we briefly review Algorithm 1. In line 1, we construct the trust and pseudo distrust relation matrix $\mathbf{F}$ and its weight matrix $\mathbf{W}$ from user-user trust relations $\mathbf{T}$, user-review authorship relations $\mathbf{P}$, and user-review helpfulness ratings $\mathbf{R}$. From line 3 to line 8, we update $\mathbf{U}$, $\mathbf{H}$ and $\mathbf{r}$ until convergence where γ_u, γ_h and γ_r are learning steps, which are chosen to satisfy Goldstein Conditions [29]. After learning the user preference matrix $\mathbf{U}$, $\mathbf{H}$ and $\mathbf{r}$ via Algorithm 1, the reconstructed trust and distrust matrix is $\hat{\mathbf{F}} = \mathbf{U}\mathbf{H}\mathbf{U}^\top + \lambda\mathbf{r}\mathbf{r}^\top$. Finally we predict pairs $\langle u_i, u_j\rangle$ whose $sign(\hat{\mathbf{F}}_{ij}) = -1$ as a distrust relation with confidence $|\hat{\mathbf{F}}_{ij}|$.

5. EXPERIMENTS

In this section, we conduct experiments to evaluate the effectiveness of the proposed framework. In particular, we try to answer two questions via experiments - (1) can the proposed framework predict distrust information indirectly with interaction data? and (2) how do the components of dTrust affect its performance? We begin by introducing experimental settings, then design experiments to seek answers for these questions and finally we do analysis on the important parameters of dTrust.

5.1 Experimental Settings

Before answering above two questions, we first introduce the experimental settings in this subsection. Let $\mathcal{A}$ be the set of pairs with trust relations in the dataset introduced in Section 3.1 and we sort $\mathcal{A}$ in a chronological order in terms of the time when pairs established trust relations. Assume that there are $x\%$ of pairs in $\mathcal{A}$ establishing trust relations until time t_x. For each x, we collect trust relations, distrust relations, user-review authorship relations and user-review helpfulness ratings until time t_x to form a evaluation dataset *Epinionsx*. In this paper, we vary x as $\{50, 70, 100\}$ and correspondingly we construct three evaluation datasets from the dataset introduced in Section 3.1, i.e., *Epinions50*, *Epinions70* and *Epinions100*. The purpose of varying the values of x is to investigate the performance of the proposed framework on Epinions datasets with different statistics.

For each dataset, we use $\mathcal{T}$ and $\mathcal{O}$ to denote sets of pairs of users with and without trust relations. $\mathcal{D}$ is the set of pairs with distrust relations, which is a subset of $\mathcal{O}$. We follow the common metric for trust/distrust evaluation in [19, 35] to assess the prediction performance. In detail, each predictor ranks pairs in $\mathcal{O}$ in a descending order of confidence and we take the first $|\mathcal{D}|$ pairs as the set of predicted distrust relations, denoting $\mathcal{P}$. Then the prediction quality is,

$$PQ = \frac{|\mathcal{P} \cap \mathcal{D}|}{|\mathcal{D}|} \tag{25}$$

where $|\cdot|$ denotes the size of a set. As noticed in [19], *the PQ value is usually low and to more meaningfully represent predictor quality, a random predictor is usually used as a baseline method*. Each experiment is repeated 10 times and we report the average performance.

5.2 Performance of Distrust Prediction

To the best of our knowledge, we are the first to study the predictability of distrust in social media; hence, there are no existing baseline methods. However, to answer the first question, we still build the following baseline methods:

- *lowTP*: Some social scientists believe distrust as the negation of trust and support that trust and distrust are two ends of the same conceptual spectrum [33, 1]. If distrust was the negation of trust, distrust information can be predicted from low trust, which can be obtained by trust prediction. *lowTP* first uses trust propagation in [10] to compute trust scores of pairs of users. The smaller the trust score, the higher the prediction confidence.

- *lowMF*: Similar to lowTP, lowMF uses low trust scores to predict distrust. lowMF uses a matrix factorization method to obtain trust scores [35].

	Epinions50	Epinions70	Epinions100
lowTP	0.00001	0.00003	0.00004
lowMF	0.00001	0.00003	0.00004
negInter	0.05458	0.11473	0.12777
random	0.00001	0.00004	0.00006
dTrust	0.08737	0.15054	0.17391

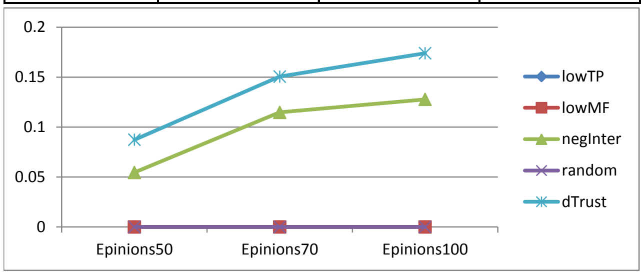


Figure 4: Performance Comparison of Different Predictors.

	Epinions50	Epinions70	Epinions100
dTrust	0.08737	0.15054	0.17391
dTrust\FW	0.00001	0.00003	0.00004
dTrust\BT	0.07336	0.13331	0.15544
dTrust\FWBT	0.00001	0.00003	0.00004
random	0.00001	0.00004	0.00006

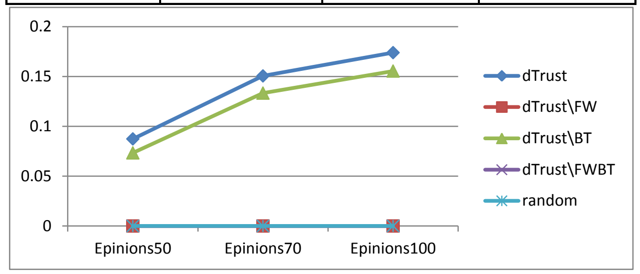


Figure 5: Component Analysis for dTrust.

- *negInter*: This method is based on the strong correlation between negative interactions and distrust relations. *negInter* ranks pairs of users based on the numbers of negative interactions. The larger the number of negative interactions, the higher the prediction confidence.
- *random*: this predictor ranks pairs of users randomly. [19] suggests that a random predictor should be used as a baseline method to meaningfully demonstrate the predictor quality since the PQ value is usually low.

For baseline methods with parameters, we try various values of these parameters and report the best performance. For dTrust, we set $\lambda = 0.1$, and empirically set $\{d = 250, \alpha = 0.1\}$. More details about parameter analysis for dTrust will be discussed later in this section. The comparison results are shown in Figure 4.

We make the following observations:

- If distrust is the negation of trust, low trust from trust prediction should accurately indicate distrust. However, we observe that most of the time, the performance of *lowTP* and *lowMF* is worse than that of *random*. These results provide evidence from the computational perspective that distrust is not the negation of trust.
- The performance of *negInter* is much better than that of *random*, which further demonstrates the existence of a strong correlation between negative interactions and distrust relations.
- dTrust always outperforms baseline methods. There are two potential contributors for this improvement. First dTrust incorporates interaction data via the trust and pseudo distrust relations $\mathbf{F}$, which is controlled by the weight matrix $\mathbf{W}$. Second, dTrust models balance theory based on trust and pseudo distrust relations from $\mathbf{F}$. More details about the effects of these components on the performance of dTrust will be discussed in the following subsection.

We perform t-test on all results and the t-test results indicate that all improvement is significant. In summary, performance comparison between the random predictor and the proposed framework dTrust suggests that dTrust can accurately predict distrust relations by incorporating interaction data and modeling balance theory, which correspondingly answers the first question.

5.3 Component Analysis for dTrust

dTrust has two important components - (1) incorporating interaction data by the user-user trust and pseudo distrust matrix $\mathbf{F}$, and (2) modeling balance theory based on $\mathbf{F}$. In this subsection, we investigate the effects of these components on the performance of dTrust to answer the second question. In detail, we systematically eliminate their effects from dTrust by defining its variants as follows,

- *dTrust*$\backslash FW$ - Eliminating the effect of incorporating interaction data $\mathbf{F}$ by defining $g(x) = 0$ in $\mathbf{W}$;
- *dTrust*$\backslash BT$ - Eliminating the effect of balance theory by setting $\lambda = 0$ in Eq. (20); and
- *dTrust*$\backslash FWBT$ - Eliminating the effects of both components by defining $g(x) = 0$ in $\mathbf{W}$ and setting $\lambda = 0$ in Eq. (20);

The results are shown in Figures 5 and "random" in the figure represents the performance of randomly guessing. The following can be observed:

- When eliminating the effect from incorporating interaction data, the performance of *dTrust*$\backslash FW$ reduces dramatically. By setting $g(x) = 0$ in $\mathbf{W}$, $\mathbf{F}$ only incorporate trust relations in $\mathbf{T}$. These results suggest not only that it is difficult, if possible, to predict distrust relations from only trust relations, but also that dTrust enables distrust prediction by incorporating interaction data.
- When eliminating the effect of balance theory, the performance of *dTrust*$\backslash BT$ degrades. Compared to dTrust, the performance of *dTrust*$\backslash BT$ reduces 0.01401, 0.01723 and 0.01847 for *Epinions50*, *Epinions70* and *Epinions100*, respectively. These results demonstrate that modeling balance theory based on $\mathbf{F}$ can improve prediction performance.

Table 2: Difference Definitions of $g(x)$ for $\mathbf{W}$. Note that "random" in the table denotes that we randomly assign values in $[0, 1]$ to the function.

	Epinions50	Epinions70	Epinions100
$g(x) = 0$	0.00001	0.00003	0.00004
$g(x) = 1$	0.05812	0.11686	0.13039
$g(x) = random$	0.05905	0.11763	0.13207
$g(x) = 1 - \frac{1}{log(x+1)}$	0.08737	0.15054	0.17391

- When eliminating both effects, the performance of $dTrust\backslash FWBT$ is the same as that of $dTrust\backslash FW$. As mentioned above, the introduction of $\mathbf{F}$ to incorporate interaction data enables the modeling of balance theory.

In summary, the introduction of the trust and pseudo distrust matrix $\mathbf{F}$ to incorporate interaction data for dTrust enables not only distrust prediction but also modeling balance theory mathematically, which can further improve the performance of dTrust.

5.4 Parameter Analysis for dTrust

There are two important parameters to control two major components of dTrust - (1) $\mathbf{W}$ controlling $\mathbf{F}$ to incorporate interaction data; and (2) λ controlling the contribution from balance theory. In this section, we investigate the impact of each of these parameters on dTrust by fixing the other to see how the performance of dTrust varies.

$\mathbf{W}$ is defined based on the function $g(x)$ to control the contribution from incorporating interaction data. We investigate the impact of $\mathbf{W}$ on the proposed framework by choosing different types of $g(x)$ for the following questions:

- What is the performance of the proposed framework if we discard the impact of the number of negative interactions by giving $g(x)$ a non-zero constant?

- If we randomly assign values of $g(x)$, what is the performance of the proposed framework dTrust?

We fix $\lambda = 0.1$, the results of dTrust with different choices of $g(x)$ are shown in Table 2. Note that "random" in the table denotes that we randomly assign values in $[0, 1]$ to the function. We make the following observations

- When $g(x) = 0$, we eliminate negative interactions and the performance reduces dramatically. This result demonstrates the importance of incorporating interaction data.

- Compared to the performance of $g(x) = 1 - \frac{1}{log(x+1)}$, the performance $g(x)$ with a non-zero constant degrades a lot. These results suggest that modeling the impact of the number of negative interactions on the correlation can improve the performance of dTrust.

- Compared to the performance of $g(x) = 1 - \frac{1}{log(x+1)}$, the performance $g(x)$ with random values also reduces a lot. These results directly suggest that $g(x)$ should not be random values, and further demonstrate the importance of modeling the impact of the number of negative interactions by $\mathbf{W}$.

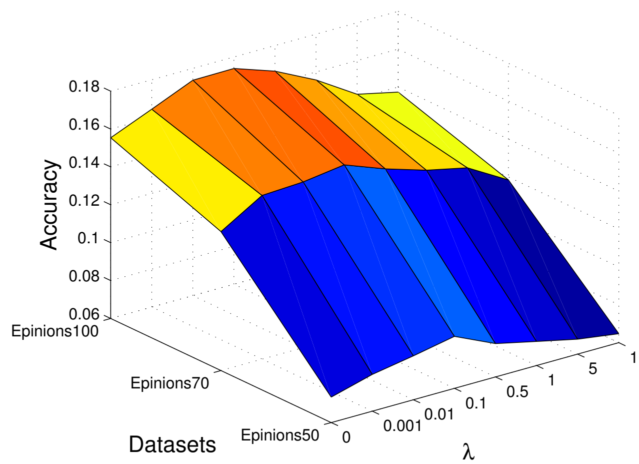


Figure 6: Impact of Balance Theory on dTrust.

After answering above two questions, we can conclude that (1) $g(x)$ in $\mathbf{W}$ should not be random values; (2) defining $g(x)$ based on the number of negative interactions can significantly improve the performance of dTrust.

To investigate the parameter λ, we set $g(x) = 1 - \frac{1}{log(x+1)}$ and vary λ as $\{0, 0.001, 0.01, 0.1, 0.5, 1, 5, 10\}$. The performance variation with respect to λ is depicted in Figure 6. In general, with the increase of λ, the performance first increases, reaches its peak value and then decreases dramatically. This pattern may be useful for us to determine the optimal value of λ in practice. In detail, when λ increases from 0 to 10, we have the following observations

- When $\lambda = 0$, eliminating the effect of balance theory from dTrust, to $\lambda = 0.1$, the performance increases a lot. For example, the accuracy increases from 0.0734 with $\lambda = 0$ to 0.0874 with $\lambda = 0.1$ in *Epinions50*. These results directly demonstrate the importance of modeling balance theory based on $\mathbf{F}$ for dTrust.

- When λ increases from 1 to 10, the performance reduces dramatically. When λ is large, the balance theory component will dominate the learning process, which may result in inaccurate estimations of $\mathbf{U}$, $\mathbf{H}$ and $\mathbf{r}$ for dTrust.

An appropriate incorporation of balance theory into dTrust based on the trust and pseudo distrust relation matrix $\mathbf{F}$ can greatly improve the performance of distrust prediction.

6. RELATED WORK

In this section, we first briefly review distrust in social sciences. As mentioned in the problem statement section, the studied problem is related to traditional trust prediction and prediction with both trust and distrust relations. Some work considers trust relations as positive relations and distrust relations as negative relations, and then trust and distrust prediction problem is converted into link prediction in signed networks [17, 42]. Therefore, we also briefly review related work from trust prediction, and link prediction in signed networks.

6.1 Distrust in Social Sciences

In social sciences, the conceptual counterpart of trust, distrust, is considered as important and complex as trust [26, 16, 11, 6]. For example, [34, 4] claim that trust and distrust help a decision maker reduce uncertainty and vulnerability (i.e., risk) associated with decision consequences; and [6] indicates that only distrust can irrevocably exclude services from being selected at all. There is a basic problem about distrust - what is the relation between trust and distrust. Answering this question is significant. If trust and distrust are the same, lack of distrust research matters little; however, if they are different, the lack of distrust research could be problematic because distrust may have unique impact. Some researchers believe distrust simply means a low level of trust, hence evidence of high trust was always regarded as being that of low distrust, and outcomes of high trust would be identical to those of low distrust [33, 1, 14]. Others believe distrust is a concept entirely separate from trust [18, 15]. For example, in [18, 27], three reasons are proposed to prove that trust and distrust are separate:(1) they separate empirically; (2) they coexist; and (3) they have different antecedents and consequents.

6.2 Trust Prediction

Most existing trust prediction algorithms can be roughly categorized into two groups - supervised methods [20, 23] and unsupervised methods [10, 35]. There are usually two steps for supervised methods. First, they extract features from available sources to represent each pair of users and consider the existence of trust relations as labels. Second, they train a binary classifier based on the representation with extracted features and labels. For example, in [20], a taxonomy is developed to systematically organize an extensive set of features for predicting trust relations and the features include user and interaction factors. User factors contain rater-related, writer-related, or commenter-related; and Viet-An Nguyen et al. [28] proposes various trust prediction models based on a well-studied Trust Antecedent Framework used in management science, capturing the three following factors: ability, benevolence and integrity. Unsupervised methods usually take advantage of some properties of trust to infer unknown trust relations. A trust propagation method is proposed in [10], which introduces four types of atomic propagations such as direct propagation, co-citation propagation, transpose propagation and trust coupling propagation. In [8], algorithms for inferring binary and continuous trust values from trust networks are proposed based on various properties of trust such as transitivity, composability and asymmetry. The user-user trust relation matrix should be low-rank and based on this property, a trust prediction algorithm is proposed based on low-rank matrix factorization in [35].

6.3 Link Prediction in Signed Networks

Link prediction in signed networks has attracted increasing attention in recent years. In [10], an algorithm based on trust and distrust propagation is proposed to predict trust and distrust relations. In [17], local-topology-based features based on balance theory are extracted to improve the performance of a logistic regression classifier in signed relation prediction. In [7], a trust and distrust prediction algorithm is proposed by combining an inference algorithm that relies on a probabilistic interpretation of trust based on random graphs with a modified spring-embedding algorithm. In [13], a low-rank matrix factorization approach with generalized loss functions is proposed to predict trust and distrust relations. Features derived from longer cycles in signed network can be used to improve link prediction performance [3]. There is also recent work to predict the signs of links. In [42], authors proposed a framework to predict the signs of a given network. Tang et al. proposed a framework to incorporate social theories into a machine learning model and infer the signs of social relations in a target network by borrowing knowledge from a different source network [40]. In [43], the authors use the transfer learning approach to leverage sign information from an existing and mature signed network to predict signs for a newly formed signed social network. The sign prediction problem is also very different from the studied problem. The sign prediction problem predicts signs for existing relations; while our problem is to predict unknown distrust relations.

7. CONCLUSION

Distrust is considered as important as trust and the value of distrust is widely recognized by social sciences. However, distrust is rarely studied in the online world because distrust information in the online world is often unavailable to the public. In this paper, we investigate whether we can obtain distrust information indirectly by studying the problem of predictability of distrust from public interaction data. We first identify that there is a strong correlation between negative interactions and distrust relations. The more negative interactions two users have, the more likely a distrust relation between them exists. Then we model a trust and pseudo distrust relation matrix from interaction data with the correlation, which not only enables distrust prediction but also allows us to model balance theory. Finally we propose a novel framework dTrust for distrust prediction by incorporating interaction data and modeling balance theory. Experimental results on real-world data show that the proposed framework dTrust can accurately predict distrust relations with interaction data. Further experiments are conducted to understand the importance of interaction data in predicting distrust relations.

There are several interesting directions needing further investigation. First since negative interactions can be found in many social media websites which only provide positive links (such as friendships in Facebook), it would be very interesting to apply the proposed framework to predict negative links with interaction data. Second, the current framework is an unsupervised method and chooses a matrix factorization method as the basic algorithm; we would like to investigate supervised methods and other basic algorithms for this problem. Finally, distrust relations inferred by the proposed framework may benefit various applications such as recommendation, and we plan to incorporate dTrust into these applications to improve performance.

Acknowledgments

The work is, in part, supported by Army Research Office (#025071), The Office of Naval Research (N000141410095) and a research fund from Yahoo Faculty Research and Engagement Program.

8. REFERENCES

[1] B. Barber. *The logic and limits of trust.* Rutgers University Press New Brunswick, NJ, 1983.

[2] D. Cartwright and F. Harary. Structural balance: a generalization of heider's theory. *Psychological Review*, 1956.

[3] K.-Y. Chiang, N. Natarajan, A. Tewari, and I. S. Dhillon. Exploiting longer cycles for link prediction in signed networks. In *CIKM*, 2011.

[4] J. Cho. The mechanism of trust and distrust formation and their relational outcomes. *Journal of Retailing*, 2006.

[5] J. Audun. Artificial reasoning with subjective logic. *Proceedings of the Second Australian Workshop on Commonsense Reasoning*, 2006.

[6] P. Cofta. Distrust. In *ICEC*, 2006.

[7] T. DuBois, J. Golbeck, and A. Srinivasan. Predicting trust and distrust in social networks. In *socialcom*, 2011.

[8] J. Golbeck. Generating predictive movie recommendations from trust in social networks. *Trust Management*, 2006.

[9] J. Golbeck. Trust and nuanced profile similarity in online social networks. *TWEB*, 2009.

[10] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *WWW*, 2004.

[11] R. Hardin. *Distrust: Manifestations and management.* Russell Sage Foundation, 2004.

[12] F. Heider. Attitudes and cognitive organization. *The Journal of psychology*, 1946.

[13] C.-J. Hsieh, K.-Y. Chiang, and I. S. Dhillon. Low rank modeling of signed networks. In *KDD*, 2012.

[14] A. Josang, E. Gray, and M. Kinateder. Analysing topologies of transitive trust. In *FAST2003*, 2003.

[15] R. M. Kramer. Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual review of psychology*, 1999.

[16] D. W. Larson and R. Hardin. Distrust: Prudent, if not always wise. *Distrust*, 2004.

[17] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. In *WWW*, 2010.

[18] R. J. Lewicki, D. J. McAllister, and R. J. Bies. Trust and distrust: New relationships and realities. *Academy of management Review*, 1998.

[19] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American society for information science and technology*, 2007.

[20] H. Liu, E. Lim, H. Lauw, M. Le, A. Sun, J. Srivastava, and Y. Kim. Predicting trusts among users of online communities: an epinions case study. In *EC*, 2008.

[21] Y. Lu, P. Tsaparas, A. Ntoulas, and L. Polanyi. Exploiting social context for review quality prediction. In *WWW*, 2010.

[22] H. Ma, D. Zhou, C. Liu, M. Lyu, and I. King. Recommender systems with social regularization. In *WSDM*, 2011.

[23] N. Ma, E. Lim, V. Nguyen, A. Sun, and H. Liu. Trust relationship prediction using online product review data. In *Proceeding of the 1st ACM international workshop on Complex networks meet information & knowledge management*, 2009.

[24] P. Massa and P. Avesani. Trust-aware bootstrapping of recommender systems. In *ECAI Workshop on Recommender Systems*, 2006.

[25] S. Chang, G. Qi, J. Tang, Q. Tian, Y. Rui and T. Huang. Multimedia LEGO: LEarning structured model by probabilistic loGic Ontology tree. In *ICDM*, 2013.

[26] D. H. McKnight and N. L. Chervany. Trust and distrust definitions: One bite at a time. In *Trust in Cyber-societies*, 2001.

[27] D. H. McKnight and V. Choudhury. Distrust and trust in b2c e-commerce: Do they differ? In *ICEC*, 2006.

[28] V. Nguyen, E. Lim, J. Jiang, and A. Sun. To trust or not to trust? predicting online trusts using trust antecedent framework. In *ICDM*,2009.

[29] J. Nocedal and S. Wright. *Numerical optimization.* Springer verlag, 1999.

[30] Y. Qian and S. Adali. *Extended structural balance theory for modeling trust in social networks.* PST, 2013.

[31] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *KDD*,2002.

[32] Z. Li, S. Chang, F. Liang, T. Huang, L. Cao and J. Smith. Learning Locally-Adaptive Decision Functions for Person Verification. In *CVPR*, 2013.

[33] J. B. Rotter. Interpersonal trust, trustworthiness, and gullibility. *American psychologist*, 1980.

[34] J. Singh and D. Sirdeshmukh. Agency and trust mechanisms in consumer satisfaction and loyalty judgments. *Journal of the Academy of Marketing Science*, 2000.

[35] J. Tang, H. Gao, X. Hu, and H. Liu. Exploiting homophily effect for trust prediction. In *WSDM*, 2013.

[36] J. Tang, H. Gao, H. Liu, and A. Das Sarma. eTrust: Understanding trust evolution in an online world. In *KDD*, pages 253–261. ACM, 2012.

[37] J. Tang and H. Liu. Trust in social computing. In *WWW*, 2014.

[38] J. Tang, X. Hu, and H. Liu. Is distrust the negation of trust? the value of distrust in social media. In *HT*, 2014.

[39] J. Tang, Y. Chang, and H. Liu. Mining social media with social theories: A survey. *SIGKDD Explorations*, 2014.

[40] J. Tang, T. Lou, and J. Kleinberg. Inferring social ties across heterogenous networks. In *WSDM*,2012.

[41] P. Victor, N. Verbiest, C. Cornelis, and M. D. Cock. Enhancing the trust-based recommendation process with explicit distrust. *TWEB*, 2013.

[42] S.-H. Yang, A. J. Smola, B. Long, H. Zha, and Y. Chang. Friend or frenemy?: predicting signed ties in social networks. In *SIGIR*, 2012.

[43] J. Ye, H. Cheng, Z. Zhu, and M. Chen. Predicting positive and negative links in signed social networks by transfer learning. In *WWW*,2013.