

# Multiple Random Walks for Personalized Ranking with Trust and Distrust

Dimitrios Rafailidis<sup>1</sup>(✉) and Fabio Crestani<sup>2</sup>

<sup>1</sup> Department of Computer Science, University of Mons, Mons, Belgium  
dimitrios.rafailidis@umons.ac.be

<sup>2</sup> Faculty of Informatics, Università Della Svizzera Italiana, Lugano, Switzerland  
fabio.crestani@usi.ch

**Abstract.** Social networks with trust and distrust relationships has been an emerging topic, aiming at identifying users' friends and foes when sharing information in social networks or purchasing products online. In this study we investigate how to generate accurate personalized rankings while considering both trust and distrust user relationships. This paper includes the following contributions, first we propose a social inference step of missing (indirect) trust relationships via multiple random walks, while considering users' direct trust and distrust relationships during the inference. In doing so, we can better capture the missing trust relationships between users in an enhanced signed network. Then, we introduce a regularization framework to account for (i) the *structural properties* of the enhanced graph with the inferred trust relationships, and (ii) the user's trust and distrust *personalized preferences* in the graph to produce his/her personalized ranking list. We evaluate the performance of the proposed approach on a benchmark dataset from Slashdot. Our experiments demonstrate the superiority of the proposed approach over state-of-the-art methods that also consider trust and distrust relationships in the personalized ranking task.

**Keywords:** Personalized ranking · Signed graphs · Social inference

## 1 Introduction

With the advent of social networks such as Epinions<sup>1</sup> and Slashdot<sup>2</sup>, users share various information through interactions with other users while expressing their positive and negative opinions. Based on their feedback, users establish trust and distrust relationships on each other, forming a signed graph with positive and negative links, respectively [13]. Node ranking with trust and distrust relationships has several applications including community detection [18], collaborative filtering [2], trust prediction [3], sign prediction [13] and troll detection [17], just to name a few. Popular ranking models like PageRank and HITS do not consider both the positive and negative links at the same time and thus, they do

<sup>1</sup> <http://www.epinions.com/>.

<sup>2</sup> <https://slashdot.org/>.

not perform well on the personalized ranking task in signed graphs. Recently, in an attempt to perform node ranking in graphs with trust and distrust relationships Shahriari and Jalili [13] review different ranking algorithms in signed graphs. In this study, they propose a modified PageRank algorithm to account for both the trust and distrust relationships in signed networks. They compute the PageRank values in a positive and a negative subgraph separately and then subtract negative PageRank values from positive ones to compute an aggregated PageRank value for each node. In the study reported in [17], a variant of PageRank is introduced to model the probability of trustworthiness of individual data sources as an interpretation for the underlying ranking values. Jung et al. [4] extend the Random Walk with Restart algorithm, namely Signed Random Walk with Restart, to generate personalized rankings in signed graphs. They introduce a random walker based on the balance theory that considers both the positive and negative links by changing the walker's sign when performing random walks on the signed graph and producing a user's ranking list.

The shortcomings of the above state-of-the-art methods are that if a user/node does not have enough interaction information, the aforementioned ranking methods face difficulties in calculating the probability of trustworthiness between two users. In addition, these ranking methods in signed networks mainly rely on the trust and distrust relationships between users based on direct signed relationships e.g., direct friend or foe, and depend on the number of common neighbours to capture the indirect relationships. So, provided that the social relationships are sparse, the challenge that we face is how can we perform personalized ranking in graphs with trust and distrust relationships to infer the missing (indirect) relationships between the users? Inferring social relationships of trust and distrust users is a challenging task [14]. Given explicit (direct) social relationships, the goal is to infer the indirect relationships of trust and distrust users. Trust relationships show strong transitivity, which means that inferring trust relationships can be computed in a network of trust users, mainly because if two users  $a$  and  $b$  are friends and a third user  $c$  is friend with  $a$ , then user  $c$  might be a friend of  $b$  as well. However, recent studies showed that distrust is certainly not transitive [1, 15, 16]. Therefore, distrust cannot be considered as the negative of trust when inferring users' distrust relationships. Accounting for the transitivity of trust relationships, a few prediction models have been proposed to infer the implicit trust relationships, while exploiting explicit distrust relationships in their predictions [3, 14]. Nonetheless, these models are designed to predict missing trust relationships and not to generate personalized rankings. Therefore, a pressing challenge resides on how to infer trust relationships of users with their distrust relationships to improve the accuracy of users' personalized rankings.

## 1.1 Contribution and Outline

As generating accurate personalized rankings becomes more and more important in social networks, in this study our contributions are the following:

- We introduce a social inference step via multiple random walks, aiming to solve the sparsity of social relationships and consequently better predict the

indirect relationships between the users. In our social inference step, we predict missing (indirect) trust relationships, while considering users' direct trust and distrust relationships during the inference.

- We propose a regularization framework for personalized ranking to consider both (i) the *structural properties* of the enhanced signed graph with the inferred relationships and (ii) the input query vector of a user with his/her *personalized preferences* of trust and distrust relationships in the signed network.

Our experiments on a benchmark dataset with trust and distrust relationships show that the proposed approach significantly outperforms other competitors in the personalized ranking prediction task in signed networks. The remainder of the paper is organized as follows, in Sect. 2 we formally define our problem, Sect. 3 details the proposed model, Sect. 4 presents the experimental results and finally, Sect. 5 concludes the study.

## 2 Problem Formulation

In this study, we consider a directed graph  $\mathcal{G}$  with  $n = |\mathcal{V}|$  nodes and  $i, j \in \mathcal{V}$ . Two nodes are connected with edges in the form  $(i, j) \in \mathcal{E}$ . The edges are considered directed and weighted, and in our setting we consider positive and negative weights to express trust and distrust relationships, respectively. Both positive and negative weights are stored in a  $(n \times n)$  weighting matrix  $\mathbf{W}$ . In our approach we generate two different graphs, a graph  $\mathcal{G}_+$  which contains only the positive edges and a second graph  $\mathcal{G}_-$  with the negative ones. Given  $\mathcal{E} \equiv \mathcal{E}_+ \cup \mathcal{E}_-$ , we compute two different  $(n \times n)$  weighting matrices  $\mathbf{W}_+$  and  $\mathbf{W}_-$ , corresponding to the weights of the positive  $(i, j)_+ \in \mathcal{E}_+$  and negative edges/relationships  $(i, j)_- \in \mathcal{E}_-$ . Notice that  $\forall (i, j)_- \in \mathcal{E}_-$  we set  $(\mathbf{W}_-)_{ij} = |\mathbf{W}_{ij}|$ , storing the absolute values if the weights of the edges are negative. In our setting, we consider a *query vector*  $\mathbf{y} \in \mathbb{R}^n$  for a node/user  $m \in \mathcal{V}$ , expressing his/her personalized preferences of trust and distrust in the signed graph. The query vector  $\mathbf{y}$  is formally defined as follows:

**Definition 1 (Query vector).** *A query vector  $\mathbf{y} \in \mathbb{R}^n$  of a user  $m$  expresses his/her personalized preferences of trust and distrust relationships in the signed graph, computed as follows,  $\mathbf{y}_i = 1$ , if  $i = m$  and  $\mathbf{y}_i = \mathbf{W}_{mi}$  otherwise.*

With these settings, the problem of personalized ranking with trust and distrust relationships is formally defined as follows:

**Definition 2 (Problem).** *“Given (i) the weighting matrices  $\mathbf{W}_+$  and  $\mathbf{W}_-$ , and (ii) a query vector  $\mathbf{y} \in \mathbb{R}^n$  of node/user  $m$  with his/her personal preferences of trust and distrust in the signed graph, the goal of the proposed approach is to generate an optimized ranking vector  $\mathbf{r} \in \mathbb{R}^n$  for ranking the  $n$  nodes, accounting for both the structural properties of the signed graph and the personalized preferences on trust and distrust relationships of user  $m$ .”*

In the first step of our approach we run multiple random walks on the graph  $\mathcal{G}_+$ , while considering the distrust relationships in graph  $\mathcal{G}_-$ . In doing so, we enhance and better capture the relationships between the trusted and distrusted nodes, inferring the missing (indirect) trust links between the nodes. Then, in the second step we consider the enhanced graph with the inferred trust and the direct (explicit) trust and distrust relationships and calculate an optimized ranking vector  $\mathbf{r}$  per node/user to produce the personalized ranking list.

### 3 Proposed Approach

#### 3.1 Social Inference via Multiple Random Walks

To infer the missing (indirect) trust relationships, we perform random walks on the  $n$  nodes in graph  $\mathcal{G}_+$ , by taking into account the distrust relationships in graph  $\mathcal{G}_-$  during the inference. In particular, the proposed approach runs multiple random walks on the graph  $\mathcal{G}_+$  with the trust relationships and then filters out the inferred trust relationships by considering the distrust relationships in graph  $\mathcal{G}_-$ . The main reason that we avoid to perform random walks on graph  $\mathcal{G}_-$  is that *distrust is not transitive*, as opposed to trust [3, 14–16]. Next, we present the case of performing a single random walk on graph  $\mathcal{G}_+$  and we show how to perform multiple random walks to better infer the implicit trust relationships.

**Single Random Walk.** Given a source node  $sou$  and a target node  $tar$ , with  $(sou, tar) \notin \mathcal{E}_+$ , the goal is to start a random walk from  $sou$  to reach  $tar$  to infer their trust relationships, denoted as  $(\overline{\mathbf{W}}_+)_{sou, tar}$ . We assume that the walk moves from one node to a neighbourhood node at each step, and at time  $t$  the walk has moved to node  $i$ . The walk chooses whether to move to another node with probability  $\xi_t$  or terminate the walk with probability  $1 - \xi_t$ . In the case of terminating the walk, the value  $(\overline{\mathbf{W}}_+)_{sou, tar}$  is returned only if edge  $(i, tar) \in \mathcal{E}_+$ , and 0 otherwise. The *transition probability* of moving from a current node  $i$  to another node  $j$  is calculated as follows:

$$p_+(j|i) = (\mathbf{W}_+)_{ij}/d_i$$

where  $d_i = \sum_j (W_+)_{ij}$  is the degree of  $i$ . The  $(n \times n)$  *transition matrix* of a random walk is given by

$$\mathbf{T}_+ = \mathbf{D}_+^{-1} \mathbf{W}_+$$

where  $(\mathbf{D}_+)_{ii} = d_i$  is the  $(n \times n)$  degree diagonal matrix. A vector  $\mathbf{p}_+^{(t)} \in \mathbb{R}^n$  represents the visiting distribution over all  $n$  nodes at a certain time  $t$ . With these settings, if the walk continues at the next time  $t + 1$ , the distribution vector will be updated as follows:

$$\mathbf{p}_+^{(t+1)} = \mathbf{p}_+^{(t)} \times \mathbf{T}_+$$

**Multiple Random Walks.** Instead of performing a single walk, we run multiple random walks from a source node in graphs  $\mathcal{G}_+$  and  $\mathcal{G}_-$  to better infer the missing trust relationships. The main reason that we can achieve better inference is that *multiple random walks start from the source user sou to seek more alternatives for the implicit (indirect) relationship to the target user tar*. Considering the graph  $\mathcal{G}_+$ , we define  $s$  as the total length of a single walk for which we recursively update the distribution vector  $\mathbf{p}^{(s)}$ . For a target node  $tar$  we consider all its in-linked edges, denoted by  $(\mathbf{W}_+)_{*tar}$ , that is the  $tar$ -th column vector of  $\mathbf{W}_+$ . With these settings, the returned value for a random walk terminated at time  $s$  is:

$$(\overline{\mathbf{W}}_+)_{sou,tar}|s = \mathbf{p}^{(s)}(\mathbf{W}_+)_{*tar} \quad (1)$$

Theoretically, we can perform random walks with infinite lengths from the source node. Aggregating the multiple random walks from the source node we have:

$$(\overline{\mathbf{W}}_+)_{sou,tar} = \sum_{t=1}^{\infty} \omega_+(t) \mathbf{p}_+^{(0)} \mathbf{T}_+^t (\mathbf{W}_+)_{*tar} \quad (2)$$

where  $\mathbf{p}_+^{(0)}$  is the starting distribution of a walk on  $\mathcal{G}_+$  and  $\omega_+(t)$  expresses the probability that a random walk will terminate at a certain time  $t$ :

$$\omega_+(t) = p_+(s = t|\xi) = \xi_t \prod_{i=1}^{t-1} (1 - \xi_i) \quad (3)$$

Therefore, the weighting matrix  $\overline{\mathbf{W}}_+$  with the inferred trust relationships is calculated as follows:

$$\overline{\mathbf{W}}_+ = \sum_{t=1}^{\infty} \omega_+(t) \mathbf{T}_+^t \mathbf{W}_+ \quad (4)$$

In our implementation, we avoid long (infinite) walks on the graph, following the idea of the ‘‘six degrees of separation’’, that is most nodes can be reached with a six step walk length [2]. This means that if a walk has reached more than six steps, then the walk is terminated. In practice, we observed that random walks do not reach more than four steps in our experiments with  $\xi_t = 0.85$ , equal to the dampening factor of PageRank [6].

When performing multiple random walks on graph  $\mathcal{G}_+$ , the distrust relationships in graph  $\mathcal{G}_-$  are ignored. Consequently, an inferred trust relationship between a source user  $sou$  and a target user  $tar$  in  $\overline{\mathbf{W}}_+$ , might have a conflict of a distrust relationship between  $sou$  and  $tar$  in graph  $\mathcal{G}_-$ . To avoid this conflict, we recompute matrix  $\overline{\mathbf{W}}_+$  by setting  $\overline{\mathbf{W}}_+ \leftarrow 0$ , if  $(\overline{\mathbf{W}}_+)_{ij} > 0 \wedge (\mathbf{W}_-)_{ij} > 0$ ,  $\forall i, j = 1 \dots n$ . Finally, the filtered trust relationships and their positive weights are stored into the initial adjacency matrix with the trust relationships, by setting  $\mathbf{W}_+ \leftarrow \overline{\mathbf{W}}_+$ .

### 3.2 Node Ranking with Trust and Distrust Relationships

**A Regularization Framework for Node Ranking.** In a plain graph with a single type of edges that considers only positive weights, we have a weighting matrix  $\mathbf{W} > 0$  and a query vector  $\mathbf{y} \in \mathbb{R}^n$  (Definition 1), where the  $i$ -th element  $\mathbf{y}_i$  is the initial query score of node  $i$ . The goal is to find a new optimized ranking vector  $\mathbf{r} \in \mathbb{R}^n$  that is smooth and close to  $\mathbf{y}$ , formulating the following cost function  $C$  as a minimization problem:

$$\min_{\mathbf{r}} C(\mathbf{r}) = S(\mathbf{r}) + \theta \widehat{E}(\mathbf{r}; \mathbf{y}) \tag{5}$$

The first term  $S(\mathbf{r})$  is a smoothness function to consider the *structural cost* of the graph. The second term  $\widehat{E}$  is the *ranking error* between the vectors  $\mathbf{r}$  and  $\mathbf{y}$  to express how well the optimized ranking vector fits the input query vector. Parameter  $\theta$  controls the influence of the second term when minimizing  $C(\mathbf{r})$ . According to the regularization framework of [19] the smoothness function can be calculated as:

$$S(\mathbf{r}) = \mathbf{r}^\top (\mathbf{I} - \mathbf{A}) \mathbf{r}$$

where  $\mathbf{A} = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$ . The ranking error is computed as follows:

$$\widehat{E}(\mathbf{r}; \mathbf{y}) = \|\mathbf{r} - \mathbf{y}\|^2 = (\mathbf{r} - \mathbf{y})^\top (\mathbf{r} - \mathbf{y})$$

To compute the optimized ranking vector  $\mathbf{r}$ , we take the gradient of  $C(\mathbf{r})$  and set it to zero:

$$\begin{aligned} \frac{\partial(C)}{\partial \mathbf{r}} &= (\mathbf{I} - \mathbf{A}) \mathbf{r} + \theta (\mathbf{r} - \mathbf{y}) = 0 \Rightarrow \\ \mathbf{r} &= (1 - \lambda) (\mathbf{I} - \lambda \mathbf{A})^{-1} \mathbf{y} \propto (\mathbf{I} - \lambda \mathbf{A})^{-1} \mathbf{y} \end{aligned} \tag{6}$$

with  $\lambda = 1/(\theta + 1) \in (0, 1)$ .

**Ranking with Trust and Distrust Relationships.** In our setting the graph contains the initial trust and distrust relationships and the inferred trust relationships in the weighting matrices  $\mathbf{W}_+$  and  $\mathbf{W}_-$ . As in the case of the plain graphs we have to calculate an optimized ranking vector  $\mathbf{r}$ , given a query vector  $\mathbf{y}$  with the user’s personalized preferences of trust and distrust in the signed graph. In this respect, we have to reformulate the smoothness function  $S(\mathbf{r})$  in Eq. (5) when minimizing the objective function  $C(\mathbf{r})$ , considering the weighting matrices  $\mathbf{W}_+$  and  $\mathbf{W}_-$ .

We combine  $\mathbf{W}_+$  and  $\mathbf{W}_-$  into a global weighting matrix  $\mathbf{W} = \mathbf{W}_+ - \mathbf{W}_-$ . As the trust and distrust relationships are directed, for a node  $i$  we calculate the out-degree  $d_i^{out} = \sum_{j \neq i} \mathbf{W}_{ij}$  and the in-degree  $d_i^{in} = \sum_{j \neq i} \mathbf{W}_{ji}$ , as well as the  $(n \times n)$  diagonal degree matrices  $(\mathbf{D}^{out})_{ii} = d_i^{out}$  and  $(\mathbf{D}^{in})_{ii} = d_i^{in}$ .

The overall smoothness of the ranking vector is the summation of all *local variations* [19]:

$$S(\mathbf{r}) = \sum_{i \in \mathcal{V}} \|\nabla_i \mathbf{r}\|^2 \tag{7}$$

where the *local variation* at each node  $n$  in our graph with trust and distrust relationships is calculated as follows:

$$\|\nabla_i \mathbf{r}\| \sqrt{\frac{1}{2} \left[ \sum_{j \neq i} \partial \mathbf{r} / \partial(i, j) + \sum_{j \neq i} \partial \mathbf{r} / \partial(j, i) \right]} \quad (8)$$

with

$$\begin{aligned} \partial \mathbf{r} / \partial(i, j) &= \sqrt{\mathbf{W}_{ij} / d_i^{out}} \mathbf{r}_i - \sqrt{\mathbf{W}_{ij} / d_j^{in}} \mathbf{r}_j \\ \partial \mathbf{r} / \partial(j, i) &= \sqrt{\mathbf{W}_{ji} / d_i^{in}} \mathbf{r}_i - \sqrt{\mathbf{W}_{ji} / d_j^{out}} \mathbf{r}_j \end{aligned} \quad (9)$$

Provided that  $\partial \mathbf{r} / \partial(i, j) = -\partial \mathbf{r} / \partial(j, i)$ , according to Eqs. (8) and (9) it is easy to verify that the overall smoothness in Eq. (7) can be reformulated as follows:

$$\begin{aligned} S(\mathbf{r}) &= \frac{1}{2} \sum_{i=1}^n \mathbf{r}_i^2 + \sum_{j=1}^n \mathbf{r}_j^2 \sum_{i=1}^n \sum_{j=1}^n \frac{\mathbf{W}_{ij} \mathbf{r}_i^2 \mathbf{r}_j^2}{\sqrt{d_i^{out} d_j^{in}}} \\ &= \mathbf{r}^\top \mathbf{r} - \mathbf{r}^\top \mathbf{D}^{out-1/2} \mathbf{W} \mathbf{D}^{in-1/2} \mathbf{r} = \mathbf{r}^\top (\mathbf{I} - \mathbf{B}) \mathbf{f} \end{aligned} \quad (10)$$

with  $\mathbf{B} = \mathbf{D}^{out-1/2} \mathbf{W} \mathbf{D}^{in-1/2}$ . Based on the formulation of the ranking problem in plain graphs in Eq. (5) and the smoothness function in Eq. (10), the ranking problem in our graph with trust and distrust relationships becomes:

$$\min_{\mathbf{r}} C(\mathbf{r}) = \mathbf{r}^\top (\mathbf{I} - \mathbf{B}) \mathbf{f} + \theta (\mathbf{r} - \mathbf{y})^\top (\mathbf{r} - \mathbf{y}) \quad (11)$$

Similar to the case of a single graph, we derive the following closed-form solution of the optimized ranking vector:

$$\mathbf{r} = (\mathbf{I} - \lambda \mathbf{B})^{-1} \mathbf{y} \quad (12)$$

To generate a *personalized ranking* for a node  $m$ , we set the query vector  $\mathbf{y}_i = 1$ , if  $i = m$  and  $\mathbf{y}_i = \mathbf{W}_{mi}$  otherwise, expressing user's  $m$  personalized preferences of trust and distrust in the signed graph (Definition 1). Having computed the query vector  $\mathbf{y}$  for the node  $m$  then we perform the personalized ranking based on Eq. (12).

## 4 Experimental Evaluation

### 4.1 Experimental Setup

As there is no groundtruth available to evaluate directly the performance of the ranking models in graphs with trust and distrust relationships, we examine the ranking performance on the troll detection task [4]. Trolls are users that can intentionally post misleading information, either having malicious intent, profit motives, or simply behaving in a disruptive way [17]. The goal of the troll detection

task is to identify trolls in a user’s personalized ranking list. In our experiments we used the “Slashdot Zoo” dataset<sup>3</sup> [5], which consists of 77,985 users, 388,190 friend (trust) links and 121,967 foe (distrust) links. Following [5], we use the foes of a user, called No-More Trolls in the “Slashdot Zoo” dataset. As we investigate the case of personalized ranking in signed graphs, we generate a personalized distrust ranking list, aiming to detect trolls high at a user’s ranking list.

In our experiments we used the ranking-based metrics precision, recall and Normalized Discounted Cumulative Gain (NDCG). Precision is defined as the ratio of the relevant items in the top- $N$  list, and recall is defined as the ratio of the relevant items in the top- $N$  ranked list over all the relevant items. The NDCG metric considers the ranking of the relevant items in the top- $N$  list. For each user the Discounted Cumulative Gain (DCG) is defined as:

$$DCG@N = \sum_{j=1}^N \frac{2^{rel_j} - 1}{\log_2 j + 1}$$

where  $rel_j$  represents the relevance score of item  $j$ , that is binary relevance in our case. As we focus on the troll detection task, we consider an item as relevant if a user is a troll, and irrelevant otherwise. NDCG is the ratio of DCG over the ideal iDCG value for each user, that is the DCG value given all trolls in the users’ personalized list. We report precision, recall and NDCG at the top- $N = 100$  results of the user’s ranking list. The reason that we consider the top-100 ranked results is that in total there are 96 trolls. We repeated our results five times and in each run we averaged the evaluation metrics over all users.

## 4.2 Compared Methods

We evaluate the performance of the following methods:

- **MPR** [13]: a Modified PageRank algorithm for ranking nodes in signed graphs. MPР computes the PageRank values in the positive  $\mathcal{G}_+$  and negative graph  $\mathcal{G}_-$  separately, and then subtracts negative PageRank values from positive ones to calculate an aggregated PageRank value per node.
- **Troll-Trust** [17]: a model that performs personalized ranking with trust and distrust relationships. Troll-Trust first uses a Bernoulli distribution to characterize each user as either being trustworthy or being a troll, and then constructs a probabilistic model based on the users’ trust and distrust relationships with an iterative algorithm.
- **SRWR**<sup>4</sup> [4]: a Signed Random Walks with Restart method for personalized ranking in signed graphs. SRWR starts a signed random surfer so that she considers negative edges by changing her sign for walking. In particular, SRWR first considers the sign of the surfer either positive or negative, that is favorable or adversarial to a node respectively, and then when a random surfer

<sup>3</sup> <http://dai-labor.de/IRML/datasets>.

<sup>4</sup> <http://datalab.snu.ac.kr/srwr>.



- encounters a negative edge, she changes her sign from positive to negative, or vice versa. Otherwise, she keeps her sign.
- **MRW-TD\***: a variant of the proposed method to evaluate the effect on the ranking performance of our model when we do not perform the inference of trust relationships with the inference step of Sect. 3.1. To achieve this, in the MRW-TD\* variant we feed the second step of our approach in Sect. 3.2 with the initial (direct) trust and distrust relationships of the original graph in the weighting matrices  $\mathbf{W}_+$  and  $\mathbf{W}_-$ , respectively.
  - **MRW-TD**: the proposed method of Multiple Random Walks with Trust and Distrust relationships.

### 4.3 Balancing Personalized Preferences of Trust and Distrust with Graphs’ Structural Properties

Figure 1 shows the effect on NDCG, when varying the  $\theta$  parameter of Eq. (11) in the proposed MRW-TD approach and its MRW-TD\* variant. While considering the structural properties of the signed graph, higher  $\theta$  values indicate that the personalized preferences of trust and distrust will influence more the objective function in Eq. (11). The results in Fig. 1 demonstrate that a selection of  $\theta = 1e-2$  achieves the best NDCG value for both methods. Setting  $\theta = 1e-1$  results in the model’s overfitting, degrading the NDCG metric for both methods. On the other hand, lower values  $\theta \leq 1e-3$  consider less user’s preferences of trust and distrust in the signed network, thus reducing the NDCG metric. Compared to the variant MRW-TD\*, MRW-TD achieves a relative improvement of 35.54%. This indicates the importance of the proposed social inference step of missing trust relationships in Sect. 3.1, which solves the sparsity in user’s social relationships. As a consequence, MRW-TD can better capture the missing relationships than its MRW-TD\* variant, hence MRW-TD produces more accurate personalized rankings, expressed by the higher NDCG values.

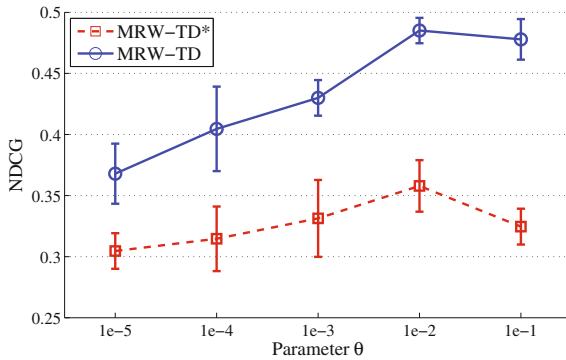


Fig. 1. Effect on NDCG when varying parameter  $\theta$ .

**Table 1.** Methods comparison in terms of NDCG, precision and recall. Bold values denote the best scores ( $p < 0.05$ ).

	NDCG	Precision	Recall
MPR [13]	.2963 $\pm$ .0243	.0812 $\pm$ .0131	.0775 $\pm$ .0107
Troll-Trust [17]	.3646 $\pm$ .0058	.1146 $\pm$ .0129	.0904 $\pm$ .0096
SRWR [4]	.4165 $\pm$ .0155	.1492 $\pm$ .0087	.1363 $\pm$ .0055
MRW-TD*	.3579 $\pm$ .0211	.1046 $\pm$ .0115	.0805 $\pm$ .0088
MRW-TD	<b>.4851 <math>\pm</math> .0104</b>	<b>.1608 <math>\pm</math> .0062</b>	<b>.1588 <math>\pm</math> .0047</b>

#### 4.4 Comparison with State-of-the-Art

In Table 1, we report average NDCG, precision and recall, to compare the proposed MRW-TD method with the state-of-the-art methods for personalized ranking in graphs with trust and distrust relationships. Compared to the second best method of SRWR, in this set of experiments we achieve a relative improvement of 16.47, 7.76 and 16.51% in terms of NDCG, precision and recall, respectively. Using the paired t-test we found that MRW-TD outperforms its competitors in all runs, with the results being statistically significant at  $p < 0.05$ . This occurs because MRW-TD performs the social inference step, and as a consequence can better capture the missing (indirect) relationships than other methods. At the same time MRW-TD balances the structural properties of the signed graph with the user’s preferences of trust and distrust in the ranking regularization framework of Eq. (11). Notice that the competitors face difficulties in the presence of sparsity in the social relationships. For instance, the second best method SRWR changes the sign of the walker based on users’ direct trust and distrust relationships and then generates the personalized rankings accordingly. However, SRWR does not predict the missing (indirect) trust relationships when producing the personalized ranking lists. Instead, our proposed MRW-TD method infers the missing trust relationships, while considering users’ direct trust and distrust relationships. Clearly, the competitors do not capture well the missing relationships, which negatively affects their ranking performance. In our approach, it is the combination of the two steps of (i) social inference of missing relationships and (ii) node ranking in the regularization framework of Eq. (11) that makes MRW-TD significantly outperform the baseline signed ranking techniques, by balancing well the structural properties of the enhanced graph with the personalized preferences of users’ trust and distrust relationships in the graph.

## 5 Conclusions

We presented an accurate personalized ranking method in signed graphs with trust and distrust relationships. As users’ social relationships are sparse, in the first step of our approach we infer missing trust relationships, while considering users’ explicit trust and distrust relationships during the inference. In addition,

we introduce a ranking regularization framework to balance users' personalized preferences of trust and distrust with the structural properties of the signed graph. Our experiments show that the proposed approach wins all the competitors by correctly inferring the missing relationships, and taking into account the graph's structural properties and user's social preferences.

Recently, collaborative ranking has gained much attention for generating personalized recommendations with trust relationships [7,8]. However, the distrust relationships are not considered in these studies. As future work we plan to extend our approach for designing a collaborative ranking model with both trust and distrust relationships, while considering evolving users' social relationships. This is a challenging task for online social networks, as users' preferences evolve over time as well [9–12].

**Acknowledgments.** Dimitrios Rafailidis was supported by the COMPLEXYS and INFORTECH Research Institutes of University of Mons.

## References

1. Guha, R.V., Kumar, R., Raghavan, P., Tomkins, A.: Propagation of trust and distrust. In: Proceedings of the 13th ACM International Conference on World Wide Web, New York, NY, USA, pp. 403–412 (2004)
2. Jamali, M., Ester, M.: Trustwalker: a random walk model for combining trust-based and item-based recommendation. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, pp. 397–406 (2009)
3. Jang, M., Faloutsos, C., Kim, S., Kang, U., Ha, J.: PIN-TRUST: fast trust propagation exploiting positive, implicit, and negative information. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, Indianapolis, IN, USA, pp. 629–638 (2016)
4. Jung, J., Jin, W., Sael, L., Kang, U.: Personalized ranking in signed networks using signed random walk with restart. In: Proceedings of 16th IEEE International Conference on Data Mining, Barcelona, Spain, pp. 973–978 (2016)
5. Kunegis, J., Lommatzsch, A., Bauckhage, C.: The slashdot zoo: mining a social network with negative edges. In: Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, pp. 741–750 (2009)
6. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: Bringing order to the web. Technical report, Stanford Digital Libraries SIDL-WP-1999-0120 (1999)
7. Rafailidis, D., Crestani, F.: Collaborative ranking with social relationships for top-n recommendations. In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, Pisa, Italy, pp. 785–788 (2016)
8. Rafailidis, D., Crestani, F.: Joint collaborative ranking with social relationships in top-n recommendation. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, Indianapolis, IN, USA, pp. 1393–1402 (2016)
9. Rafailidis, D., Kefalas, P., Manolopoulos, Y.: Preference dynamics with multimodal user-item interactions in social media recommendation. *Expert Syst. Appl.* **74**, 11–18 (2017)

10. Rafailidis, D., Nanopoulos, A.: Modeling the dynamics of user preferences in coupled tensor factorization. In: Proceedings of the 8th ACM Conference on Recommender Systems, Foster City, Silicon Valley, CA, USA, pp. 321–324 (2014)
11. Rafailidis, D., Nanopoulos, A.: Repeat consumption recommendation based on users preference dynamics and side information. In: Proceedings of the 24th International Conference on World Wide Web Companion, Florence, Italy, pp. 99–100 (2015)
12. Rafailidis, D., Nanopoulos, A.: Modeling users preference dynamics and side information in recommender systems. *IEEE Trans. Syst. Man Cybern.: Syst.* **46**(6), 782–792 (2016)
13. Shahriari, M., Jalili, M.: Ranking nodes in signed social networks. *Social Netw. Analys. Min.* **4**(1), 172 (2014)
14. Tang, J., Chang, Y., Aggarwal, C., Liu, H.: A survey of signed network mining in social media. *ACM Comput. Surv.* **49**(3), 42:1–42:37 (2016)
15. Tang, J., Hu, X., Chang, Y., Liu, H.: Predictability of distrust with interaction data. In: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, Shanghai, China, pp. 181–190 (2014)
16. Victor, P., Cornelis, C., Cock, M.D., Teredesai, A.: Trust- and distrust-based recommendations for controversial reviews. *IEEE Intell. Syst.* **26**(1), 48–55 (2011)
17. Wu, Z., Aggarwal, C.C., Sun, J.: The troll-trust model for ranking in signed networks. In: Proceedings of the 9th ACM International Conference on Web Search and Data Mining, San Francisco, CA, USA, pp. 447–456 (2016)
18. Zhang, J., Zhan, Q., He, L., Aggarwal, C.C., Yu, P.S.: Trust hole identification in signed networks. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, Riva del Garda, Italy, pp. 697–713 (2016)
19. Zhou, D., Schlkopf, B.: A regularization framework for learning from graph data. In: ICML Workshop on Statistical Relational Learning and Its Connections to Other Fields, pp. 132–137 (2004)