Table 1: Statistics of User-Item Rating Matrix of Epinions

Statistics	User	Item
Max. Num. of Ratings	1960	7082
Avg. Num. of Ratings	12.21	7.56

Table 2: Statistics of Social Trust Network of Epinions

Statistics	Trust per User	Be Trusted per User
Max. Num.	1763	2443
Avg. Num.	9.91	9.91

the product or service with an integer rating value from one to five. Each member of Epinions maintains a trust list which presents a network of trust relationships among users.

The Epinions dataset consists of 51,670 users who have rated a total of 83,509 items. The total number of ratings is 631,064, and the total number of the issued trust statements is 511,799. Note that the user-item rating matrix of Epinions dataset is quite sparse, and its density is around 0.015%. The statistics of user-item matrix and user-trust matrix are summarized into Table 1 and Table 2;

The different amounts of training data (80% and 90%) are used to evaluate the algorithms. The training data 80% means that we randomly choose 80% of ratings from the Epinions dataset as the training data to predict the ratings of the remaining 20%. The random selection was performed five times independently. For the proposed method, the regularization coefficient λ is empirically set to one, and η is 0.1. All of our experiments were carried on a PC with Intel 2.8GHz processor and 4GB RAM.

5.2 Performance Measure

Both the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are employed as the performance measure to evaluate the prediction quality of our proposed approaches in comparison with other related recommendation methods. More specifically, the metrics RMSE is defined as follows:

$$\textbf{RMSE} = \sqrt{\frac{1}{p}\sum_{i,j}(M_{i,j} - \hat{M}_{i,j})^2}$$

where $M_{i,j}$ denotes the rating on the item j annotated by the user i, and $\hat{M}_{i,j}$ is the predicted rating for the user i on the item j. p is denoted as the total number of ratings in the testing dataset. Additionally, the metrics MAE is defined as below:

$$\textbf{MAE} = \frac{1}{p}\sum_{i,j}|M_{i,j} - \hat{M}_{i,j}|$$

To make a consistent comparison with the previous methods, we employ the inverse mapping function of $g(x)$ in Eqn. 2 to transform the predicted rating value in our method back into the original range.

5.3 Experiments on Epinions Dataset

The rank d is an important parameter for the matrix factorization problem. Instead of assigning it to an empirical value like the previous methods, we can estimate d based on our LRSDP formulation. For the proposed approach, we observe that the number of variables is $d(m+n)$ and the total number of equality constraints is equal to the number of ratings p. To guarantee a well-posed problem, the rank d should be less than $\frac{p}{m+n}$. Moreover, we did not take into account of both the inactive users and those items having no ratings in practice. Therefore, the rank d is set to five.

In this paper, we employ the user-trust information from Epinions dataset to build the graph Laplacian matrix L. As discussed in Section 4.1, we have three choices on selecting the kernel function for the graph Laplacian. Table 3 presents the experimental results of different settings. From the empirical results, we first observe that both the distance kernel and heat kernel outperform the binary kernel, which indicates that it is effective to incorporate the users' taste by comparing their ratings. Second, the distance kernel function using the modified cosine similarity achieves the best performance in the empirical evaluation. Therefore, we choose the modified cosine distance function to build the graph Laplacian matrix L in the following experiments.

Table 3: Evaluation on the different kernel functions (A Smaller RMSE or MAE Value Means a Better Performance)

Training Data	Metrics	Binary	Distance	Heat
80%	RMSE	1.1379	**1.1304**	1.1323
	MAE	0.8639	**0.8557**	0.8606
90%	RMSE	1.1155	**1.1095**	1.1197
	MAE	0.8540	**0.8338**	0.8457

Then, we conduct the empirical comparisons to investigate the effectiveness of the logistic function $\rho(x)$ for normalization. Moreover, we study the LRSDP method without the graph Laplacian regularization, which is equivalent to the low-rank semidefinite program solution for the minimization problem in Eqn. 3. To make it clear, the proposed LRSDP with the graph Laplacian regularization approach is denoted as "LRSDP". The LRSDP method without normalization is denoted as "LRSDP(w/o ρ)". Similarly, the LRSDP method without the graph Laplacian regularization is denoted as "LRSDP(w/o L)". Table 4 shows the experimental results for the proposed methods and the state-of-the-art approaches. We can observe that LRSDP(w/o ρ) obtains the very poor results, which reveals the significance of normalization using the logistic function. Moreover, the proposed LRSDP approach outperforms the two recent social recommendation approaches: STE and SoRec. This demonstrates that the graph Laplacian regularization not only stands on a solid theoretical framework but also obtains the promising results in practice. Looking into the performance comparisons, we can also find that LRSDP(w/o L) performs slightly better than its counterpart PMF method using gradient descent optimization.

Finally, we empirically study the efficiency performance of the proposed LRSDP method. Table 5 summarize the computational time for factorizing the user-item rating matrix using STE and LRSDP. From these results, it can be clearly observed that the proposed LRSDP approach is much