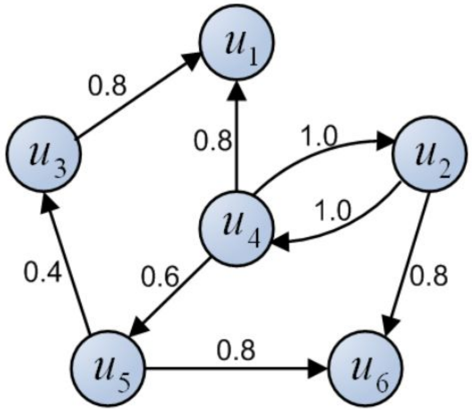
Trust Aware Recommendation Systems

Problem



(a) Social Network Graph

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2		3		4		
u_2	4	3			5			
u_3	4		2				2	4
u_4								
u_5	5	1	2		4	3		
u_6	4	3		2	4		3	5

(b) User-Item Matrix

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2	**2.5**	3	**4.8**	4	**2.2**	**4.8**
u_2	4	3	**2.4**	**2.9**	5	**4.1**	**2.6**	**4.7**
u_3	4	**1.7**	2	**3.2**	**3.9**	**3.0**	2	4
u_4	**4.8**	**2.1**	**2.7**	**2.6**	**4.7**	**3.8**	**2.4**	**4.9**
u_5	5	1	2	**3.4**	4	3	**1.5**	**4.6**
u_6	4	3	**2.9**	2	4	**3.4**	3	5

(c) Predicted User-Item Matrix

Regularization Methods : SocialMF

Regularization methods focus on a user’s preferences and assume that a user’s preferences should be similar to that of her trust network.

SocialMF

$$\min \sum_{i=1}^{n} (\mathbf{u}_i - \sum_{u_k \in \mathcal{N}_i} \mathbf{T}_{ik} \mathbf{u}_k)^2,$$

$$\min_{\mathbf{U},\mathbf{V}} \|\mathbf{W} \odot (\mathbf{R} - \mathbf{U}^\top \mathbf{V})\|_F^2 + \alpha \sum_{i=1}^{n} (\mathbf{u}_i - \sum_{u_k \in \mathcal{N}_i} \mathbf{T}_{ik} \mathbf{u}_k)^2$$
$$+ \lambda(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$$

SocialMF

$$\mathcal{L}(R,T,U,V) = \frac{1}{2}\sum_{u=1}^{N}\sum_{i=1}^{M} I_{u,i}^{R}(R_{u,i} - g(U_u^T V_i))^2$$

$$+ \frac{\lambda_U}{2}\sum_{u=1}^{N} U_u^T U_u + \frac{\lambda_V}{2}\sum_{i=1}^{M} V_i^T V_i$$

$$+ \frac{\lambda_T}{2}\sum_{u=1}^{N}\Big((U_u - \sum_{v\in N_u} T_{u,v}U_v)^T(U_u - \sum_{v\in N_u} T_{u,v}U_v)\Big)$$

SocialMF

$$\frac{\partial \mathcal{L}}{\partial U_u} = \sum_{i=1}^{M} I_{u,i}^{R} V_i g'(U_u^T V_i)(g(U_u^T V_i) - R_{u,i}) + \lambda_U U_u$$
$$+\lambda_T (U_u - \sum_{v \in N_u} T_{u,v} U_v)) - \lambda_T \sum_{\{v|u \in N_v\}} T_{v,u} \Big(U_v - \sum_{w \in N_v} T_{v,w} U_w \Big) \quad (13)$$

$$\frac{\partial \mathcal{L}}{\partial V_i} = \sum_{u=1}^{N} I_{u,i}^{R} U_v g'(U_u^T V_i)(g(U_u^T V_i) - R_{u,i}) + \lambda_V V_i \quad (14)$$

Efficiency

Users/Items	Root Mean Square Error	RMSE in paper	Training Data
7000/21000	1.12	1.075	90%
3000/9000	1.078	1.075	90%
7000/21000	1.17	1.075	80%
3000/9000	1.13	1.075	80%

Regularization Methods: SocialMF

One advantage of these approaches is that they indirectly model the propagation of tastes in social networks, which can be used to mitigate cold-start problem and increase the coverage of items for recommendations.

Co-factorization Methods

- The underlying assumption of systems in this group is that the i-th user u_i should share the same user preference vector u_i in the rating space (rating information) and the trust relation space.
- Systems in this group perform a co-factorization in the user-item matrix and the user-user trust relation matrix by sharing the same user preference latent factor.

Sorec

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{F}_i} (\mathbf{S}_{ik} - \mathbf{U}_i^\top \mathbf{Z}_k)^2,$$

$$\min_{\mathbf{U},\mathbf{V},\mathbf{Z}} \|\mathbf{W} \odot (\mathbf{R} - \mathbf{U}^\top \mathbf{V})\|_F^2 + \alpha \sum_{i=1}^{n} \sum_{u_k \in \mathcal{F}_i} (\mathbf{S}_{ik} - \mathbf{U}_i^\top \mathbf{Z}_k)^2$$
$$+ \lambda(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{Z}\|_F^2),$$

Hybrid: Sorec and SocialMF

- The i-th user u_i should share the same user preference vector u_i in the rating space (rating information) and the trust relation space.
- A user’s preferences should be similar to that of his trust network.

Sorec

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{F}_i} (\mathbf{S}_{ik} - \mathbf{U}_i^\top \mathbf{Z}_k)^2,$$

$$\min \sum_{i=1}^{n} (\mathbf{u}_i - \sum_{u_k \in \mathcal{N}_i} \mathbf{T}_{ik} \mathbf{u}_k)^2,$$

Has Potential

- 90% Train, 10% Test, 3000 Users, 9000 items.

SocialMF : 1.078 RMSE

Hybrid : 1.054 RMSE

- This was the least error across multiple runs.

Zomato

- Very few datasets with trust relations and item ratings.
- Most just use epinions, some papers introduce new ones with slightly different density.
- Api exists but only for restaurants and their ratings.
- Idea was to use “followers” as one way trust relation.

Zomato: Challenges

- Default scraping didn’t work.
- User agent spoofing, but got static web pages.
- Selenium worked but very slow.
- Had to find post URL and use it for scraping.

Zomato: Comparison with Epinions

Parameters\Datasets	Epinions	Zomato
Users	49,290	8028
Items	1,39,738	92,324
Items Ratings	6,64,823	2,33,974
Trust Relations	4,87,182	38,84,886

More Trust, Lesser Error

- 90% Train, 10% Test, 3000 Users, 9000 items. Sorec.

Epinions : 0.87 MAE

Zomato : 0.53 MAE

Overview

- Identified core ideas/papers in Trust Based Recommendation Systems and implemented and compared them to know advantages and disadvantages of each.
- Tried to leverage previous understanding to create a hybrid approach and improve performance.
- Contributions: Created a new dataset with high trust density.

Thanks!

Student Info:

P. Rishith Reddy
201401159

rishith.reddy@students.iiit.ac.in