Recommendation in Social Networks

Martin Ester

Simon Fraser University

Tutorial at RecSys 2013

Outline

- Introduction
- Recommendation in social networks
- Memory based approaches for item recommendation
- Model based approaches for item recommendation
- Top-N item recommendation
- Friend recommendation
- Future research

Introduction

- Users want to have personalized results.
- But are not willing to spend a lot of time to specify their personal information needs.
- Recommender systems automatically identify information relevant for a given user, learning from available data.
- Data
  - user actions,
  - user profiles.

Introduction

Rating Matrix

Items

Target Item

Users		Departed	Star Wars	Matrix	Hurt Locker	Titanic	Terminator
Target User	Joe	2	5	4	2	?	?
	John	5	1	2		1	
	Susan	5			5	5	5
	Pal	2	5		3		
	Jean		5	3	5		3
	Ben		1			5	
Similar User	Nathan	2		4	1	4	

Ratings

Introduction

- Rating prediction

  Predict the rating of target user for target item, e.g. predict Joe's rating for Titanic.
- Top-N item recommendation

  Predict the top-N highest-rated items among the items not yet rated by target user.
- Friend recommendation (only if social network) Predict the top-N users to which the target user is most likely to connect.

Social Networks

[Wasserman & Faust 1994]

- Used widely in the social and behavioral sciences, in economics, marketing, . . .
- Directed or undirected graph

  nodes: actors
  edges: social relationships or interactions

Social Networks

- Different types of social relationships
- Different types of interactions

Social Networks

- Explicit social network
  relationships provided by users

- Implicit social network
  relationships inferred from user actions
  - Email network
  - Co-worker network

Social Networks

[Monge & Contractor 2003]

- The formation and evolution of social networks is affected by many effects, including
  - Self-interest,
  - Social and resource exchange,
  - Balance,
  - Homophily,
  - Proximity.

Trust Networks

[Golbeck 2005]

- Trust network allows users to
  - systematically document their trust-relationships,
  - see which users have declared trust in another user.
- Connected users do not necessarily have a social relationship.
- Trust in a user may be based, e.g., on articles or reviews authored by that user.
- Trusted users influence other users.

Online Social Networks

- Emergence of online social networks
- Among the top websites http://www.alexa.com/topsites

. . .

2. FaceBook

. . .

8. LinkedIn

. . .

10. Twitter

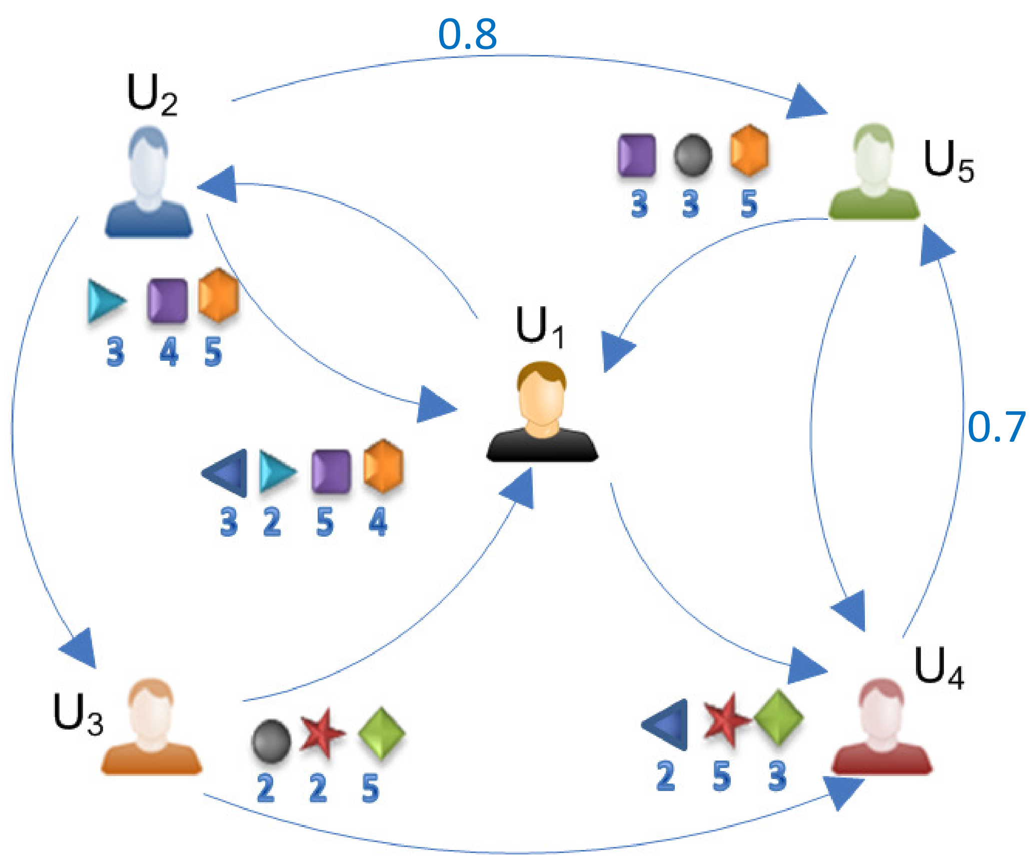
. . .

→ Availability of very large datasets

Social Rating Networks

- Social rating network (SRN):
  social network, where users are associated with item ratings.
- Item ratings can be numeric [1..5]
  or Boolean (bookmark photo, like article, . . .).
- Examples: Epinions, Flixster, last.fm,
  flickr, Digg.
- Social action: create social relationship,
  rating action: rate an item.

Social Rating Networks



Effects in Social Rating Networks

- Social influence:
  ratings are influenced by ratings of friends,
  i.e. friends are more likely to have similar ratings than strangers.
- Correlational influence:
  ratings are influenced by ratings of actors with similar ratings,
  i.e. if some ratings are similar, further ratings are more likely also to be similar.

Effects in Social Rating Networks

- Selection (homophily):
  actors relate to actors with similar ratings,
  i.e. actors with similar ratings are more likely to become friends.
- Transitivity:
  actors relate to friends of their friends,
  i.e. actors are more likely to relate to indirect friends.

Recommendation in Social Networks

SFU

- Benefits of social network-based recommendation:
  - Exploit social influence, correlational influence, transitivity, selection.
  - Can deal with cold-start users, as long as they are connected to the social network.
  - Are more robust to fraud, in particular to profile attacks.

Recommendation in Social Networks

SFU

- Challenges
  - Low probability of finding rater of target item at small network distance.
  - Ratings at large network distances are noisy.
  - Social network data is sensitive (privacy concerns).
  - Edges in online social networks are of greatly varying reliability / strength.

Influence and Correlation

[Anagnostopoulos et al. 2008]

- Question: does a SN exhibit social influence?
- Discrete time period [0..T], consider only one action, e.g. using a certain tag.
- At every time step, each user flips a coin to decide whether he will get active.
- Probability of activation depends only on number *a* of already active friends:

$$p(a) = \frac{e^{\alpha \ln(a+1)+\beta}}{1+e^{\alpha \ln(a+1)+\beta}}$$

Influence and Correlation

SFU

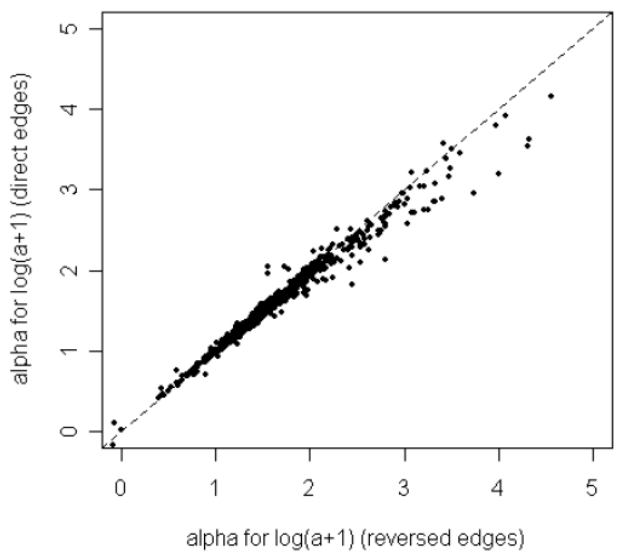
- α measures social correlation
- $Y_{a,t}$: number of users with a active friends at time t-1 who get activated at time t
- $N_{a,t}$: number of users with a active friends at time t-1 who do not get activated at time t
- $Y_a = \sum_t Y_{a,t}$, $N_a = \sum_t N_{a,t}$
- Compute α and β that maximize the data likelihood $\prod_a p(a)^{Y_a} (1-(p(a))^{N_a}$

Influence and Correlation

- If social influence plays a role, then the timing of activation should depend on the timing of activation of other users.
- $W = \{w_1, \ldots, w_l\}$: set of active users at time T
- t_i: activation time of user *i*
- Shuffle test
  – Perform random permutation π of $\{1, \ldots, l\}$.
  – Set activation time of user *i* to $t_{i'} := t_{\pi(i)}$.

Influence and Correlation

- Compute α for original activation times.
- Compute α' for shuffled activation times.
- If α and α' are close to each other, then the model exhibits no social influence.



α for original activation times vs. α' for shuffled activation times on Flickr dataset

→ Social correlation,
but no social influence!

Inferring Social Relationships and Their Strength

- Often: (Boolean) social network given.
- Sometimes, no information about social relationships, only user actions.

  → Inference of social network from user actions
- Often, users have many friends, and the strength of the relationships varies greatly.

  → Inference of weighted social network

Inferring Social Networks

[Gomez-Rodriguez et al. 2010]

- Goal: infer social relationships from user actions with time stamps.
- Assumption: there is a latent, static network over which influence propagates.
- t_u: activation time of user *u*, i.e. time when user *u* gets activated ("infected") by a cascade
- Cascade *c* specified through activation times of all users: $c = [t_1, \ldots, t_n]$, possibly $t_i = \infty$

Inferring Social Networks

- Independent Cascade model:

  activated node activates each of his friends with a given probability

- $P_c(u,v)$: probability of cascade *c* spreading from user *u* to user *v*
- $\Delta = t_v - t_u$
- $P_c(u,v)$ decreases with increasing Δ

$$P_c(u,v) \propto e^{-\frac{\Delta}{\alpha}} \text{ or } P_c(u,v) \propto \frac{1}{\Delta^{\alpha}}$$

Inferring Social Networks

- C: set of all given cascades
- G: inferred directed graph over node (user) set U={1,…,n}
- T(G): set of all subtrees of G

$$P(c\,|\,T) = \prod_{(u,v)\in T} P_c(u,v)$$

$$P(C\,|\,G) = \prod_{c\in C} \max_{T\in T(G)} P(c\,|\,T) = \prod_{c\in C} \max_{T\in T(G)} \prod_{(u,v)\in T} P_c(u,v)$$

- Problem: compute *G* with at most *k* edges that maximizes the likelihood $P(C\,|\,G)$

Inferring Social Networks

- Improvement of log-likelihood over empty graph *E*: $F_c(G) = \max_{T \in T(G)} \log P(c \mid T) - \max_{T \in T(E)} \log P(c \mid T)$
- Equivalent problem

$$G^* = \underset{|G| \leq k}{\operatorname{argmax}} F_C(G)$$

- Problem is NP-hard.
- $F_C(G)$ is submodular, which means that a greedy algorithm gives a constant-factor approximation of the optimal solution.
  → NetInf algorithm

Inferring Weighted Social Networks

[Myers et al. 2010]

- NetInf is very accurate for homogeneous networks, i.e. networks where all connected nodes influence ("infect") each other with the same probability.
- For inhomogeneous networks, define

$$A_{ij} = P(\text{node } i \text{ infects node } j \mid \text{node } i \text{ is infected})$$

- Goal: learn the matrix A = [A_{ij}] from the observed set of cascades C={c_1,. . . , c_n}

Inferring Weighted Social Networks

- If *i* becomes infected, then *j* will be infected with probability A_{ij}.
- w(*t*): transmission time model
  probability distribution of the transmission time from one node to a friend

τ_i^c : time of infection of node *i* by cascade *c*

time of infection of *i*'s friend *j* by cascade *c*

$$\tau_j^c = \tau_i^c + t, \text{where } t \sim w(t)$$

Inferring Weighted Social Networks

SFU

- Likelihood of observed cascades C given a weight matrix A

$$P(C \mid A) = \prod_{c \in C} \left[\left(\prod_{i:\tau_i^c < \infty} \left(1 - \prod_{j:\tau_j^c < \tau_i^c} (1 - w(\tau_i^c - \tau_j^c) A_{ji}) \right) \right) \left(\prod_{i:\tau_i^c = \infty} \prod_{j:\tau_j^c < \infty} (1 - A_{ji}) \right) \right]$$

- First term: one factor for each infected node *i*, assuming that at least one of his friends *j* who was infected earlier infected him.
- Second term: one factor for each non-infected node *i*, assuming that none of the infected friends *j* infected him.

Data Sets for Recommendation in SNs

- Epinions
  - Online product reviews.
  - Explicit notion of trust.
  - Users review and rate products in different categories.
  - Users express trust on other reviewers.
  - http://www.trustlet.org/wiki/Epinions_dataset
    - 50K users, 140K items, 650K ratings, 480K links
  - http://alchemy.cs.washington.edu/data/epinions/
    - 70K users, 105K items, 575K ratings, 500K links
  - 50 % cold start
    - Less than 5 ratings

Data Sets for Recommendation in SNs

- Flixster
  - Social networking service for rating movies.
  - Friendship relations.
  - http://www.sfu.ca/~sja25/datasets/
  - 1M users, 50K items, 8M ratings, 26M links
  - 85% of users have no ratings
  - 50% of raters are cold start
    less than 5 ratings

Approaches for Recommendation in SNs

- Memory based approaches
  - Explore the social network for raters.
  - Aggregate the ratings to compute prediction.
  - Store the social rating network.
  - No learning phase.
  - Slow in prediction.
  - First generation of recommenders in SN were memory based approaches.

Approaches for Recommendation in SNs

- Model based approaches
  - Learn a model.
  - Store the model parameters only.
  - Substantial time for learning.
  - Fast in prediction.
  - Most methods are based on matrix factorization.

Memory based Approaches

- Explore the network to find raters in the neighborhood of the target user.
- Aggregate the ratings of these raters to predict the rating of the target user.
- Different methods to calculate the “trusted neighborhood” of users.

TidalTrust

[Golbeck 2005]

- Modified breadth-first search in the network.
- Consider all raters *v* at the shortest distance from target user *u*.
- Trust between *u* and *v*

$$t_{u,v} = \frac{\sum_{w \in N_u} t_{u,w} t_{w,v}}{\sum_{w \in N_u} t_{u,w}}$$

where N_u denotes set of (direct) neighbors (friends) of *u*

- Trust depends on all connecting paths.

TidalTrust

- Predicted rating

$$\hat{r}_{u,i} = \frac{\sum_{v \in raters} t_{u,v} r_{v,i}}{\sum_{v \in raters} t_{u,v}}$$

where $r_{v,i}$ denotes rating of user v for item i

- Only considers raters at the shortest distance:
  - Efficient,
  - High precision,
  - Low recall.

MoleTrust

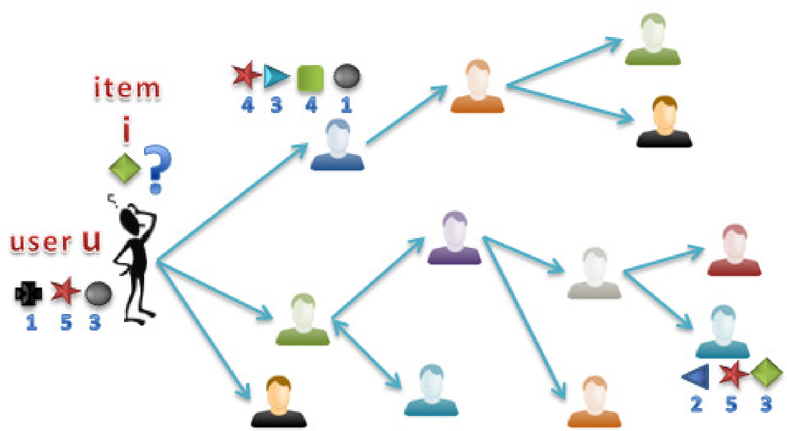
[Massa & Avesani 2007]

- Similar to the idea of TidalTrust.
- Considers raters up to a maximum-depth *d*.
- Tuning *d*:
  Trade-off between precision (and efficiency) and recall.

TrustWalker

[Jamali & Ester 2009]

- How far to explore the network?
- Instead of far neighbors who have rated the target item use near neighbors who have rated similar items.



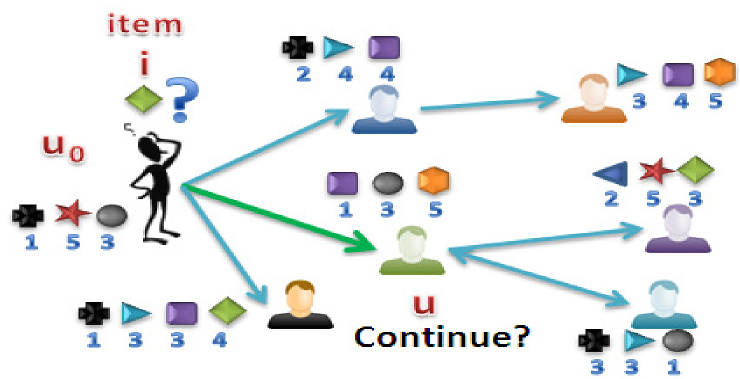
TrustWalker

- Random walk based model.
- Combines item-based recommendation and trust-based recommendation.
- Performs several random walks on the network.
- Each random walk returns a rating of the target item or a similar item.
- Prediction = aggregate of all returned ratings.

TrustWalker

SFU

- Each random walk starts from target user u_0.



- At step k, at node u:
  - If u has rated i, return $r_{u,I}$.
  - With $\Phi_{u,i,k}$, stop random walk, randomly select item j rated by u and return $r_{u,j}$.
  - With $1-\Phi_{u,i,k}$, continue the random walk to a direct neighbor of u.

TrustWalker

- Item similarities

$$sim(i,j) = \frac{1}{1+e^{-\frac{|UC_{i,j}|}{2}}} \times corr(i,j)$$

- $\Phi_{u,i,k}$ depends on
  - Similarity of items rated by *u* and target item *i*
  - And the step of random walk:

$$\phi_{u,i,k} = \max_{j \in RI_u} sim(i,j) \times \frac{1}{1+e^{-\frac{k}{2}}}$$

Experimental Evaluation

SFU

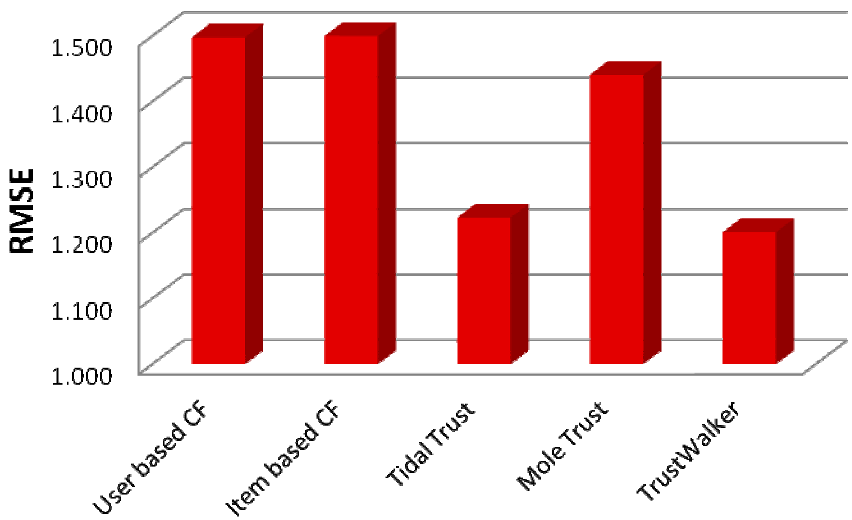
RMSE results on Epinions

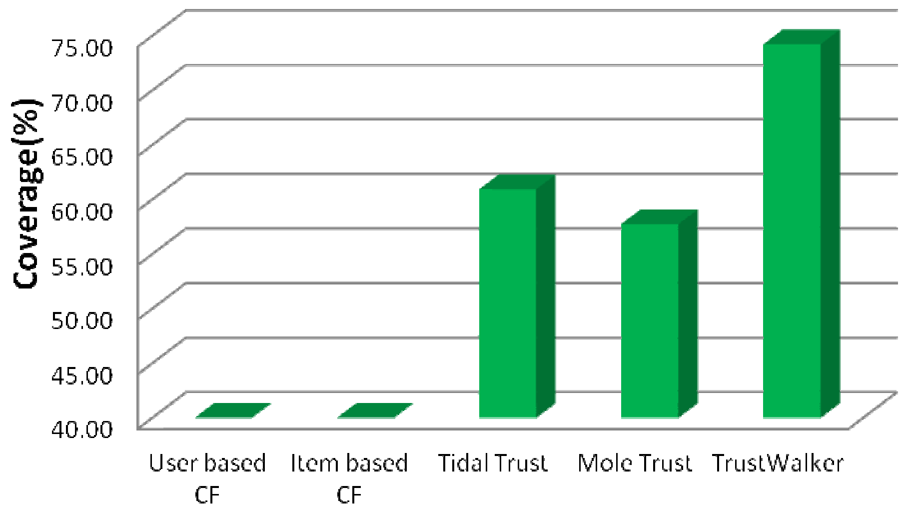
Method	Cold Start Users			All Users		
	RMSE	Coverage(%)	F-Measure	RMSE	Coverage(%)	F-Measure
Item based CF	1.551	21.26	0.316	1.232	68.91	0.691
User based CF	1.498	16.34	0.259	1.277	67.54	0.688
MoleTrust	1.441	55.36	0.594	1.104	81.03	0.765
TidalTrust	1.223	56.92	0.626	1.109	82.37	0.770
TrustWalker	1.201	70.17	0.701	1.079	93.22	0.819

RMSE results on Flixster

Method	Cold Start Users			All Users		
	RMSE	Coverage(%)	F-Measure	RMSE	Coverage(%)	F-Measure
Item based CF	1.097	71.59	0.721	0.8938	94.27	0.852
User based CF	1.114	69.86	0.710	0.9132	90.37	0.833
MoleTrust	1.083	95.93	0.829	0.8997	95.46	0.856
Tidal Trust	1.106	96.11	0.826	0.8821	96.12	0.861
TrustWalker	1.042	96.51	0.837	0.8413	99.63	0.881

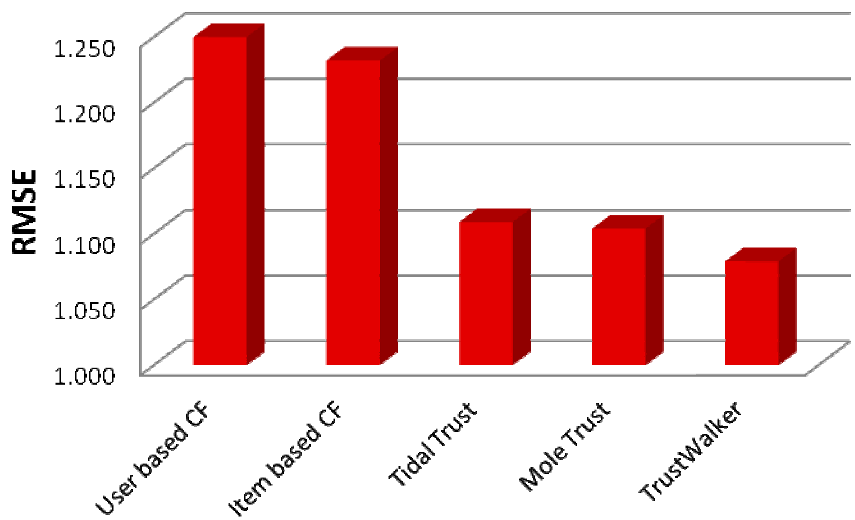
Experimental Evaluation

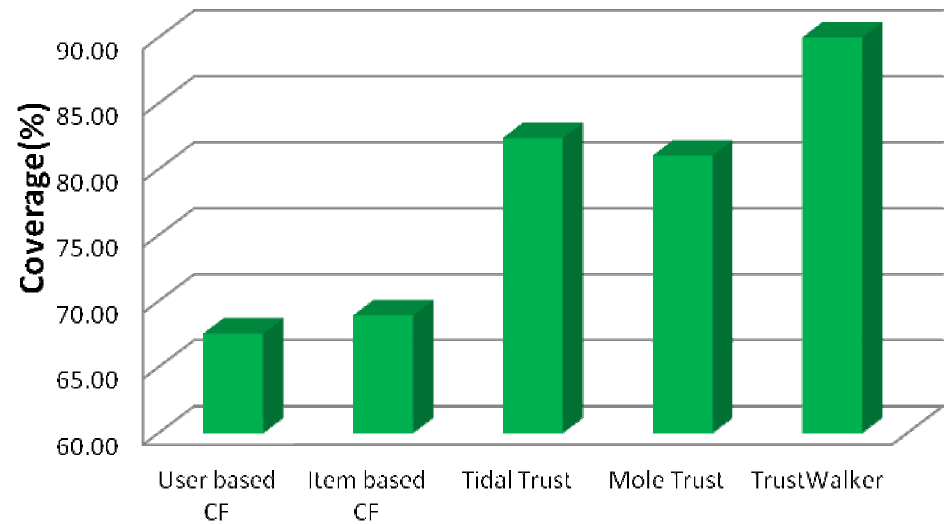




Results for cold start users on Epinions

Experimental Evaluation





Results for all users on Epinions

Model based approaches

- Matrix factorization [Koren et al. 2009]
- Observed ratings $R_{u,i}$
- Latent factors for users

$$U \in \mathbb{R}^{K \times N}$$

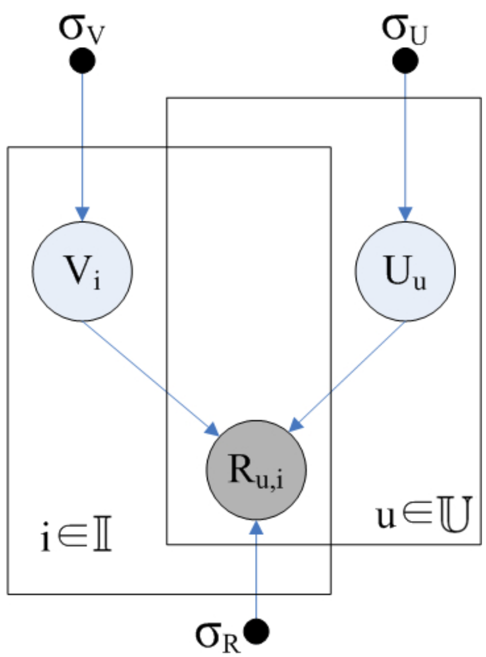
- Latent factors for items

$$V \in \mathbb{R}^{K \times M}$$

$$p(R|U,V,\sigma_R^2) = \prod_{u=1}^{N} \prod_{i=1}^{M} \left[\mathcal{N}\left(R_{u,i}|U_u^T V_i, \sigma_r^2\right)\right]^{I_{u,i}^R}$$

Model based approaches

Graphical model



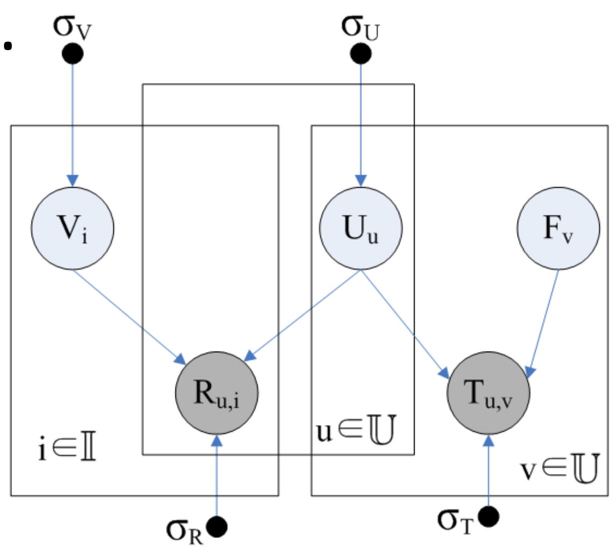
- Learn *U*, *V* that minimize

$$\sum_{all\ observed\,(u,i)} (R_{ui} - \widehat{R}_{ui})^2 + \lambda(\|U\|^2 + \|V\|^2)$$

SoRec

[Ma et al. 2008]

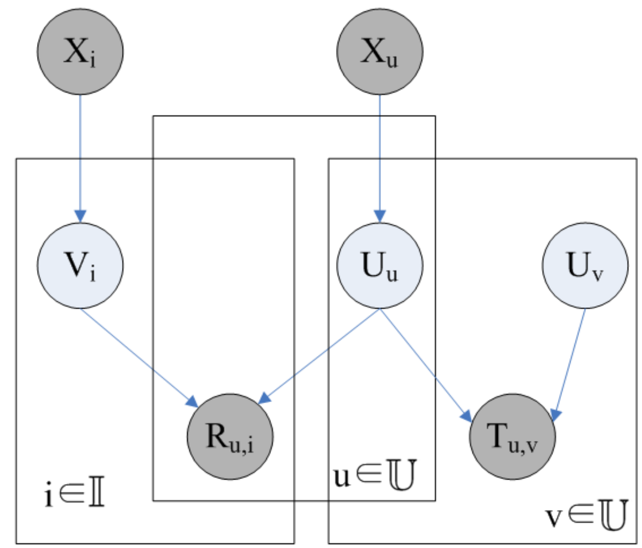
- Matrix factorization model
  - Factorize the ratings and links together.
  - Social network as a binary matrix.
- One latent factor for items.
- Two latent factors for users:
  - One for the initiator,
  - One for the receiver.
- Same user factor for both contexts (rating actions and social actions).



FIP

[Yang et al. 2011]

- Factorizes both rating matrix and the social network.
- Similar to SoRec.
- Assumes undirected network.
- FIP vs. SocRec
  - SocRec: Two user factors, FIP one user factor.
  - FIP uses user / items features as priors for user / item factors.



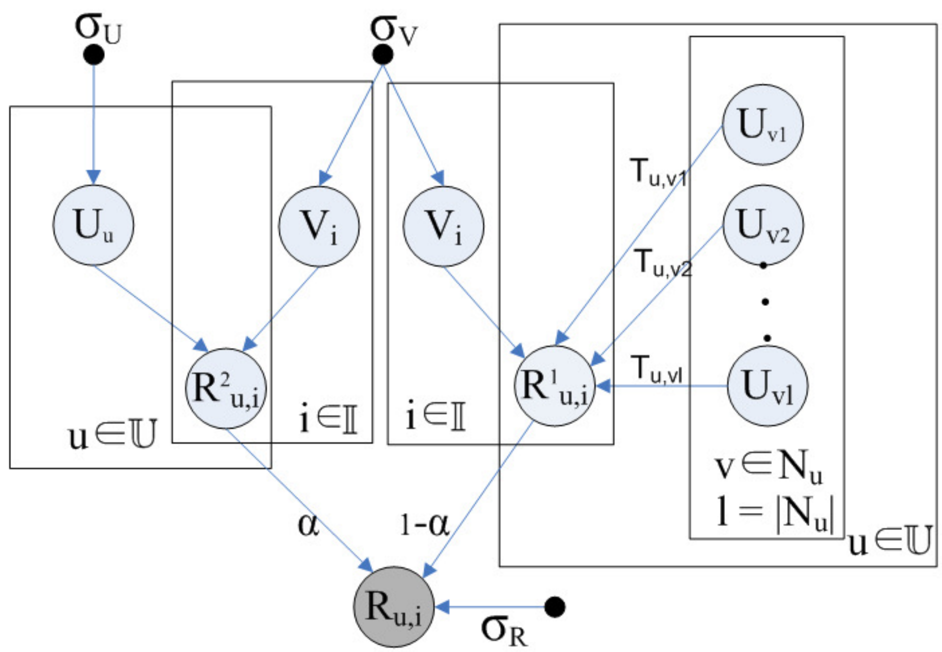
Social Trust Ensemble

[Ma et al. 2009]

- Social Trust Ensemble (STE)
- Linear combination of
  – Basic matrix factorization and
    Latent factors of the user and the item determine the observed rating.
  – Social network based approach
    Latent factors of the neighbors and the latent factor of the item determine the observed rating.

Social Trust Ensemble

- Graphical model



$$\hat{R}_{u,i} = \alpha U_u^T V_i + (1-\alpha) \sum_{v \in N_u} T_{u,v} U_v^T V_i$$

Martin Ester: Recommendation in Social Networks, Tutorial at RecSys 2013

Social Trust Ensemble

- Issues with STE
  - Learning of user factors is based on observed ratings only.
  - STE does not handle trust propagation.

SocialMF

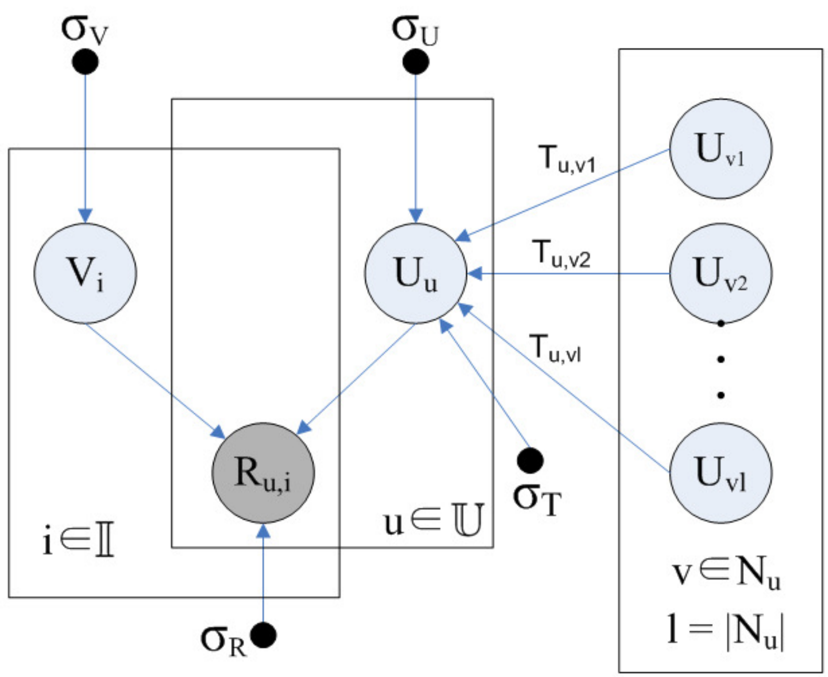
[Jamali & Ester 2010]

- Social influence: behavior of a user u *is* affected by his direct neighbors N_u.
- Latent factors of a user depend on those of his neighbors.

$$\widehat{U}_u = \sum_{v \in N_u} T_{u,v} U_v$$

- $T_{u,v}$ is the normalized trust value.

SocialMF



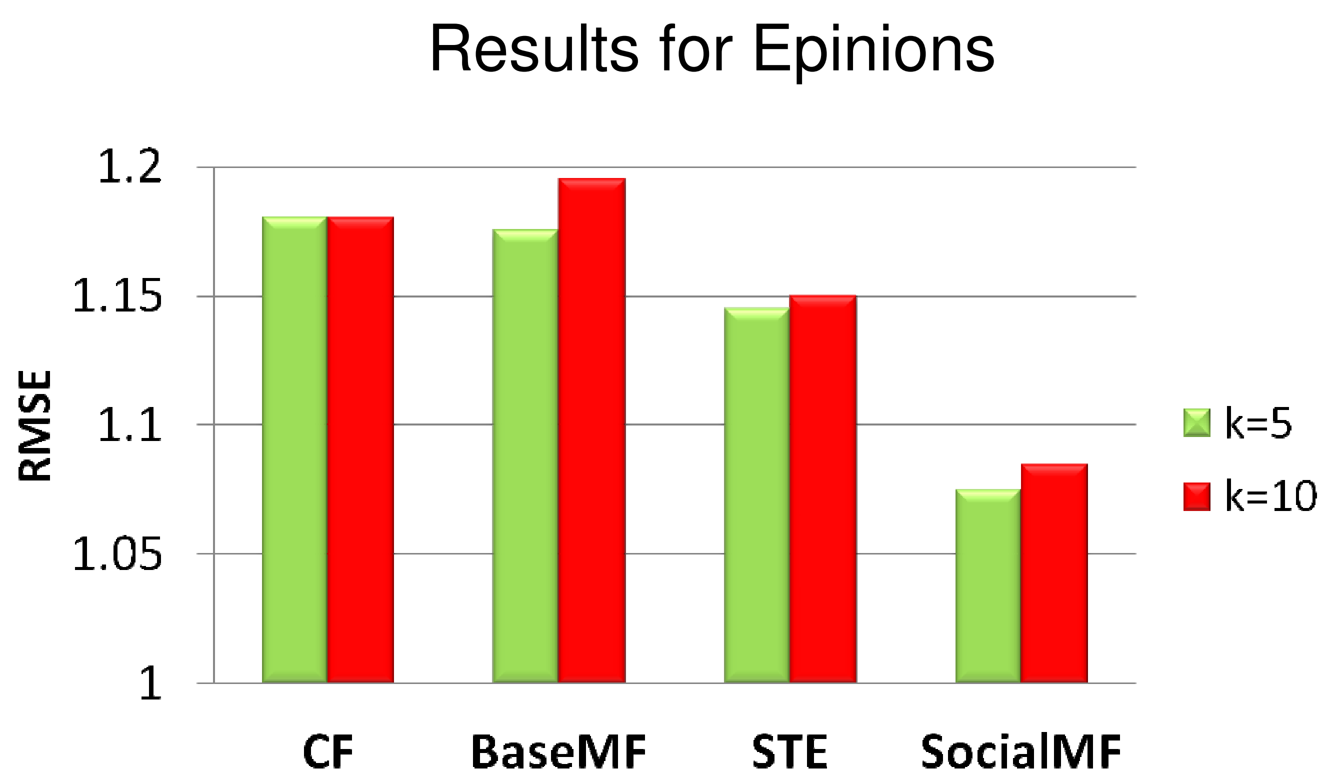
$$\sum_{all\ observed(u,i)} (R_{ui} - \widehat{R}_{ui})^2 + \lambda(\|U\|^2 + \|V\|^2)$$

$$+\beta\left(\sum_u \left(\left(U_u - \sum_v T_{u,v}U_v\right)\left(U_u - \sum_v T_{u,v}U_v\right)^T\right)\right)$$

SocialMF

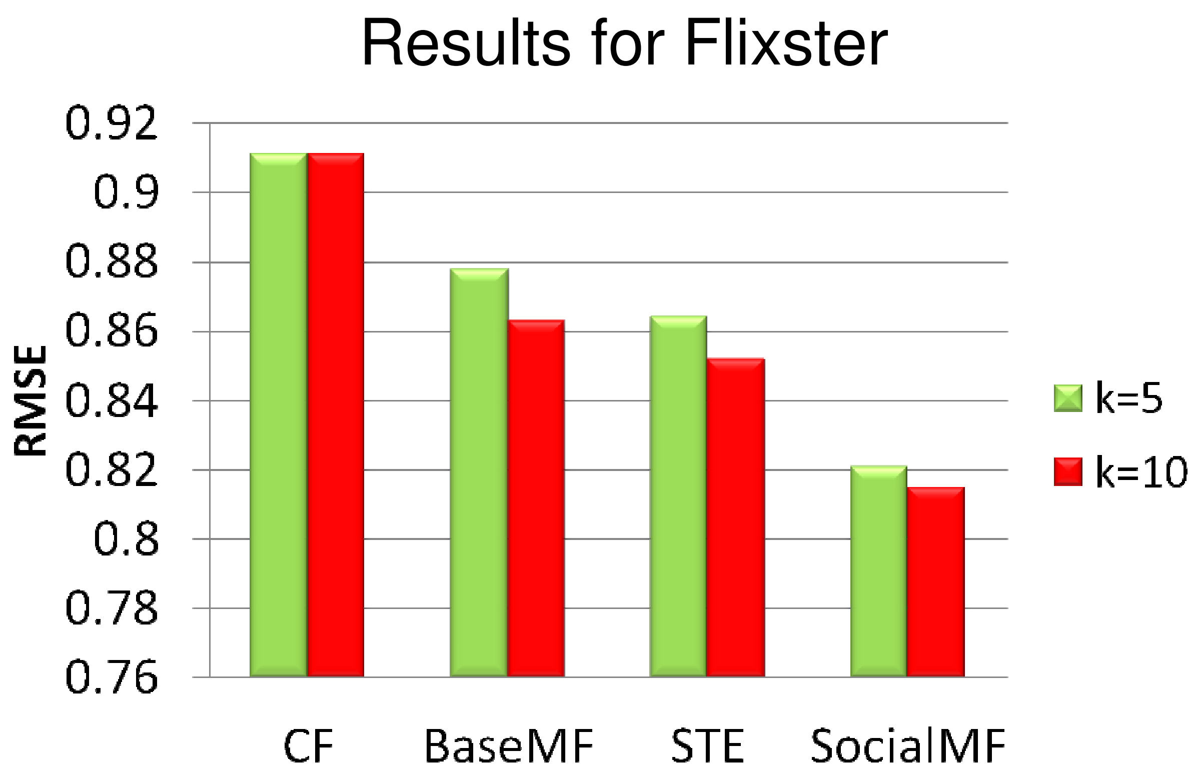
- Properties of SocialMF
  - Models trust propagation.
  - Learning the user latent factors is possible with social network only.
  - Works for cold start users and even users with no ratings.
- Similar ideas in [Ma et al. 2011].

Experimental Evaluation



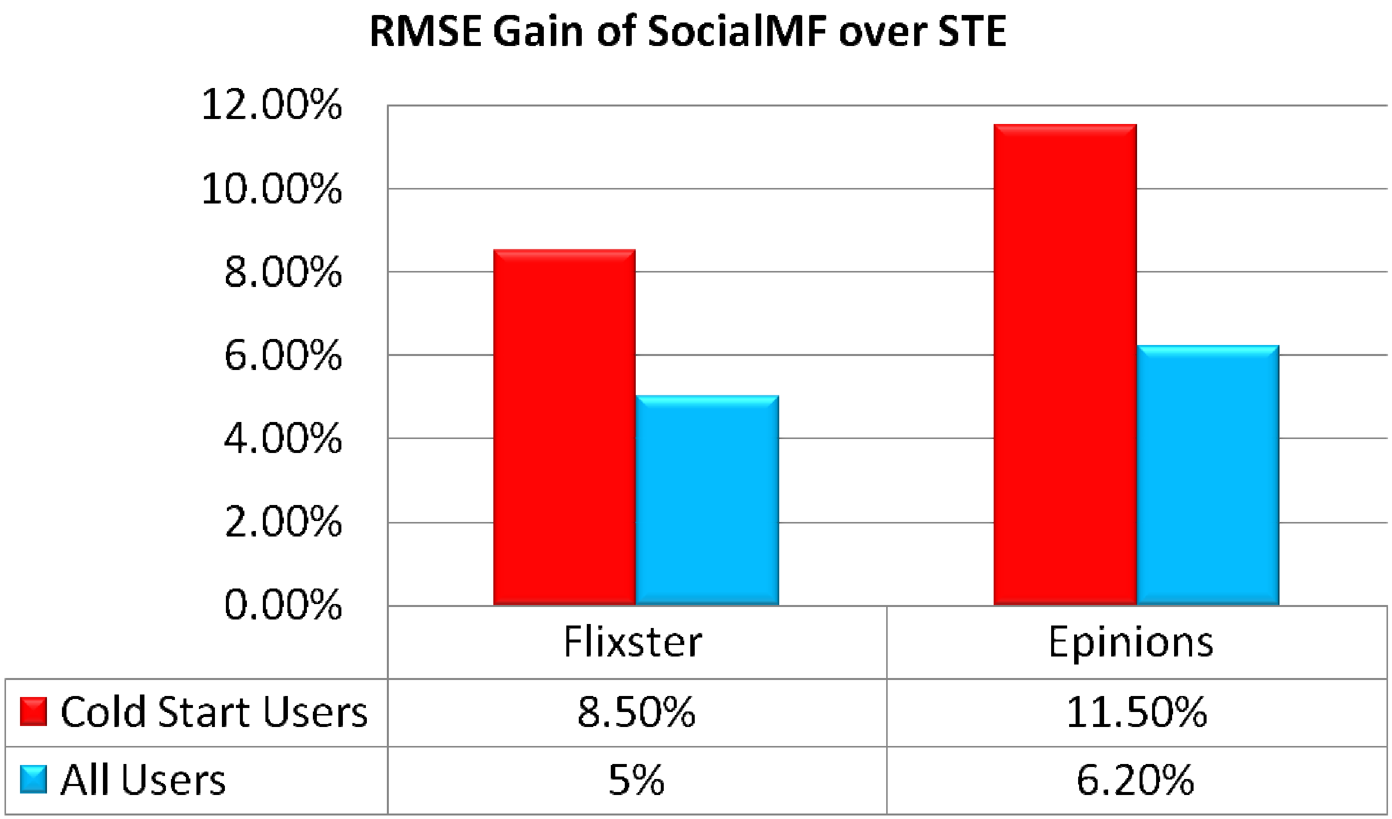
- Gain over STE: 6.2%. for K=5 and 5.7% for K=10

Experimental Evaluation



- SocialMF gain over STE (5%) is 3 times the STE gain over BasicMF (1.5%)

Experimental Evaluation



Generalized Stochastic Block Model

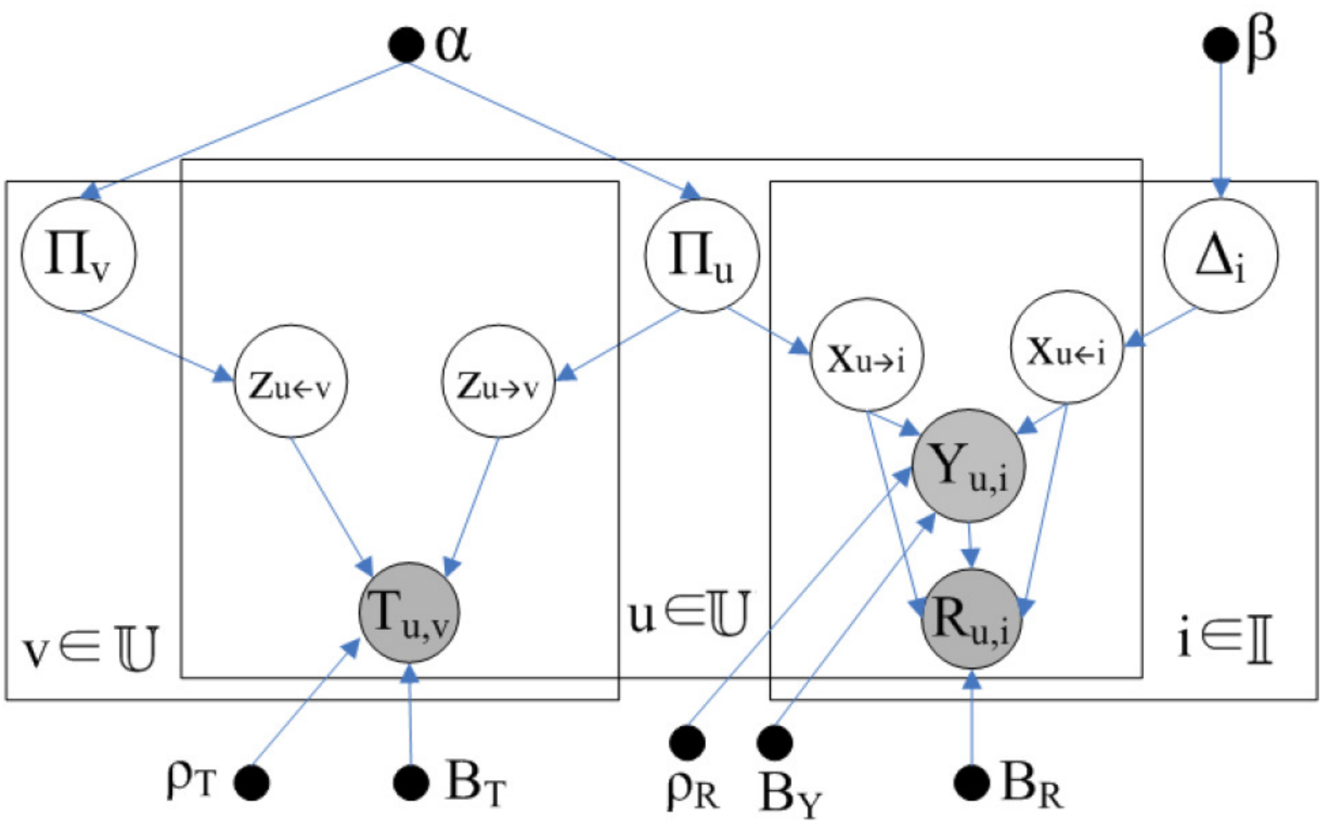
[Jamali et al. 2011]

- Social influence and selection lead to formation of communities/groups.
- Users belong to different (latent) groups,
  e.g. teacher interacting with students or his/her son or camera being rated by professional vs. amateur.
- Items belong to different (latent) groups,
  e.g. high-quality and low-quality items.
- Clustering based method for recommendation.

Generalized Stochastic Block Model

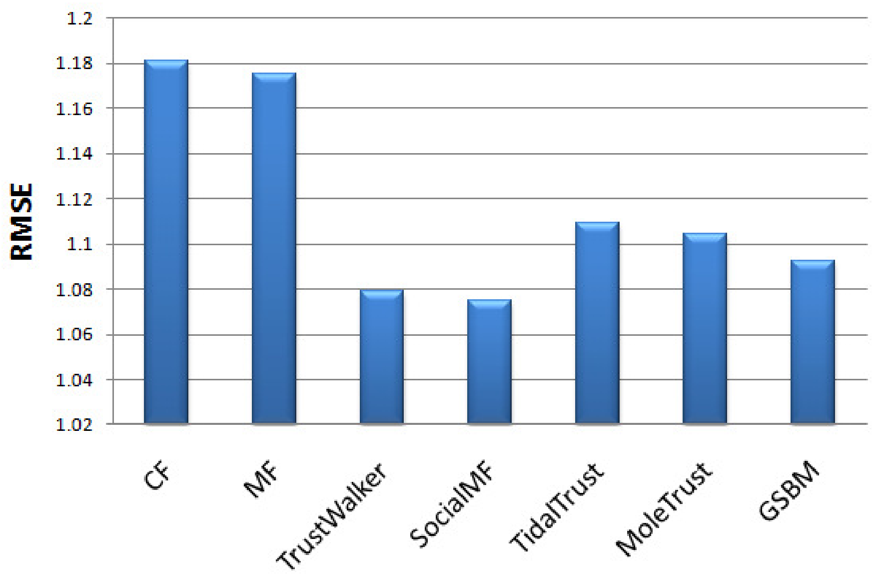
- Extending mixed membership stochastic block model [Airoldi et al. 2008].
- Users probabilistically act as a member of one of the groups in their actions.
- Every item is considered to belong to a group when it is being rated.
- The relation between users and items is governed by the relation between their groups.

Generalized Stochastic Block Model

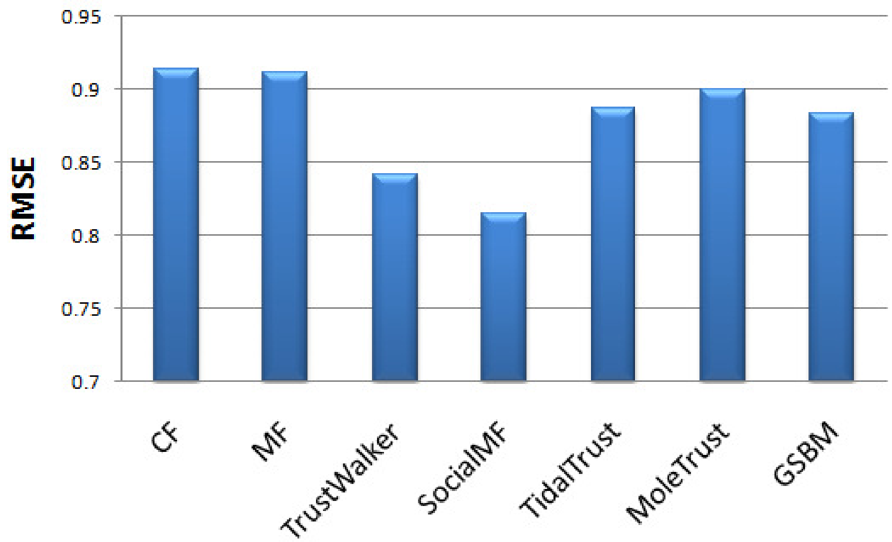


Experimental Evaluation

Results for item recommendation



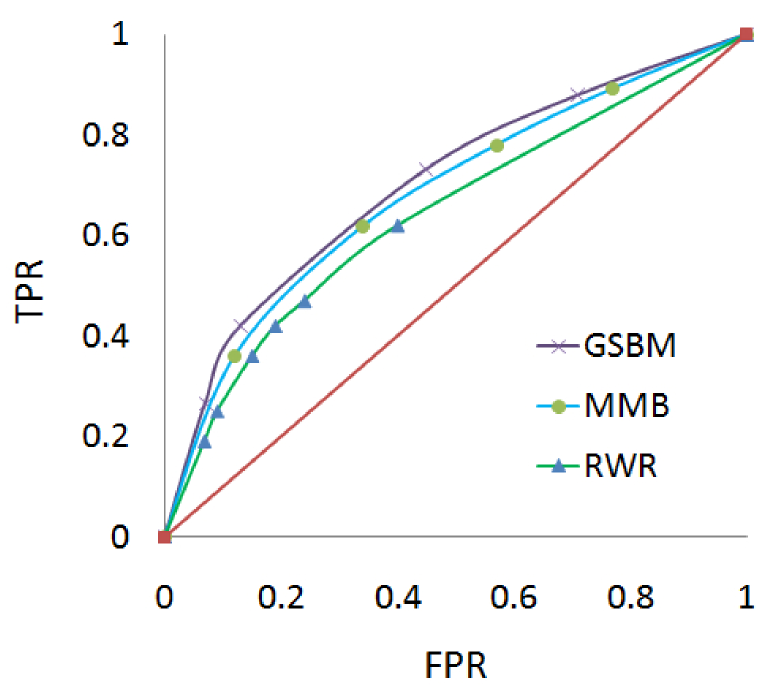
Epinions



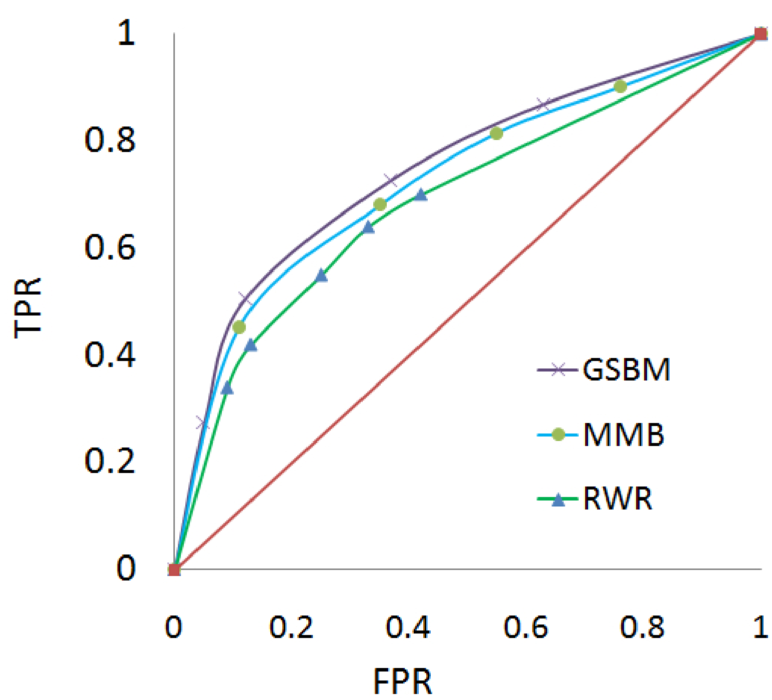
Flixster

Experimental Evaluation

Results for friend recommendation



Epinions



Flixster

Top-N Item Recommendation

- So far, we (and most of the literature) have considered only rating prediction.
- Learning objective is to minimize the prediction error computed on the observed ratings in the training set.
- However, top-N recommendation more relevant in practical applications.
- Can use rating prediction method for top-N item recommendation by ranking all items without observed rating in descending order of predicted rating.

Top-N Item Recommendation

- Better to change learning objective: minimize the ranking error for all pairs of items whose ratings are observed in the training set.
- Problem 1: This is not efficient.
- Problem 2: Ratings are missing not at random (MNAR). Low ratings are typically much more likely to be missing. [Steck 2010]
- Not a problem for rating prediction, since the data in the training set and the test set are from the same biased distribution.

Top-N Item Recommendation

[Yang et al. 2012]

- But a problem for top-N recommendation, as the *N* recommended items have to be chosen from *all* items that were not rated in the training set.
- Solution: compute rating prediction error for all items, not only for those whose ratings are observed in training set.
- Impute small constant rating value r_m for unobserved ratings.
- Give smaller weight to prediction errors for unobserved ratings.

Top-N Item Recommendation

- Can easily modify existing rating prediction methods for top-N item recommendation.
- E.g. for MF, replace term of the objective function

$$\sum_{all\ observed\,(u,i)} (R_{ui} - \hat{R}_{ui})^2 + \lambda(\|U\|^2 + \|V\|^2)$$

by

$$\sum_{all\ (u,i)} (R_{ui} - \hat{R}_{ui})^2 + \lambda(\|U\|^2 + \|V\|^2)$$

Experimental Evaluation

Results (recall top-500) on Epinions

test users	MF models			
	No Trust	SocialMF	STE	SoRec
original training (on observed ratings)				
all	1.9%	3.5%	2.7%	2.6%
cold	1.5%	1.0%	2.8%	2.9%
modified training (on all ratings)				
all	26.0%	29.1%	29.4%	32.0%
cold	16.5%	27.9%	26.6%	33.3%

- Modified objective drastically improves recall.
- Relative performance opposite of rating prediction.

Experimental Evaluation

Results (recall top-100) on Flixster

test users	MF models			
	No Trust	SocialMF	STE	SoRec
original training (on observed ratings)				
all	4.4%	4.7%	5.3%	8.2%
cold	6.3%	6.6%	7.2%	15.4%
modified training (on all ratings)				
all	44.3%	45.2%	47.1%	49.1%
cold	30.8%	38.3%	47.6%	59.2%

- Modified objective drastically improves recall.
- Relative performance opposite of rating prediction.

Friend Recommendation

Facebook friend recommendations

- „People you may know“
- “Based on mutual friends, work and education information, networks you’re part of, contacts and many other factors.”
- “Since our formula is automatic, you might occasionally see people you don’t know or don’t want to be friends with. To remove them from view, just click the X next to their names.”

http://www.facebook.com/help/?page=199421896769556

Friend Recommendation

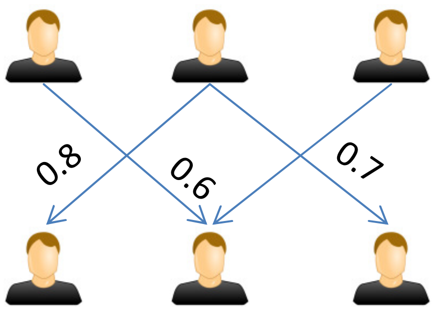
- Significance of friend recommendation
  - All major social networks have it.
  - E.g., in LinkedIn 50% of connections from recommendations. [Posse 2012]
- Problem definitions
  - Given a user pair (*u*,*v*), estimate the probability of creation of the link u→v.
  - Given a user *u*, recommend a list of top users for *u* to connect to.

Friend Recommendation

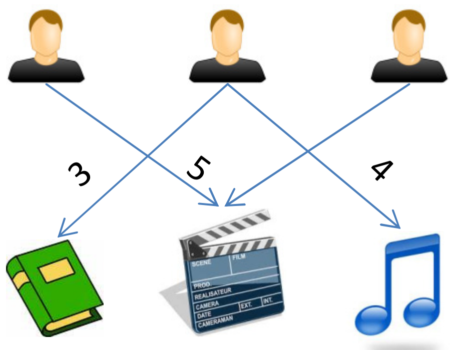
- How does it work?
- For instance, two people may meet at a party and then get connected on Facebook.
- Reasons to connect often exogeneous to the SN, but the SN contains clues.
- For instance, two people are more likely to meet at a party, if
  - They are close to each other in the SN,
  - They have similar age,
  - They live in the same town.

Friend Recommendation

- Friend recommendation vs. item recommendation



Strength of a relation between a user and another user



Strength of a relation between a user and an item

Friend Recommendation Methods

- Roots in social selection.
- Users with highest similarity to *u* are recommended to *u*.
- Every user *u* is represented by his/her observed properties such as neighbors and past activities such as ratings and clicks.

Friend Recommendation Methods

SFU

- General similarity measure between users A and B
  - the ratio between the amount of information needed to state the commonality of A and B and the information needed to fully describe what A and B are [Lin 1998]:

$$sim(A, B) = \frac{\log P(common(A, B))}{\log P(description(A, B))}$$

- Special Cases:
  - Cosine similarity
  - Pearson correlation
  - Jaccard's coefficient

Topology based Methods

- Measure similarity, based on direct neighbors of *A* and *B*.
- Common neighbors $$score(A, B) = |N_A \cap N_B|$$
- Jaccard's coefficient $$score(A, B) = \frac{|N_A \cap N_B|}{|N_A \cup N_B|}$$
- [Adamic & Adar 2003]

$$score(A, B) = \sum_{C \in N_A \cap N_B} \frac{1}{\log |N_C|}$$

Topology based Methods

- Preferential attachment [Newman 2001]

$$score(A,B) = |N_A|.|N_B|$$

    - Initially proposed for modeling network growth.

- SimRank [Jeh & Widom 2002]
  - Two user are similar to the extent that they are joined to similar neighbors.

$$score(A,B) = \gamma \cdot \frac{\sum_{x \in N_A} \sum_{y \in N_B} score(x,y)}{|N_A| \cdot |N_B|}$$

$$score(x,x) = 1$$

Path based Methods

- Measure similarity, based on paths between A and B.
- Katz [Katz 1953]:

$$score(A,B) = \sum_{l=1}^{\infty} (\beta^l \cdot paths^l{}_{A,B})$$

  - $path^l{}_{A,B}$: number of paths of length l from *A* to *B*
- Hitting time [Liben-Nowell &Kleinberg 2003]
  - score(A,B): Average number of steps for a random walk from *A* to *B.*

Path based Methods

- Random walk with restart [Pan et al. 2004]
  - A random walk starts from A. At each step, with probability α the random walk restarts.
  - score(A,B):
    probability of being at B during random walk from A.

Path based Methods

- Supervised random walks [Backstrom & Leskovec 2011]
- Learn edge weights so that random walk visits more likely "positive" nodes (to which new edges will be created in the future) than "negative" nodes (other nodes).
- Supervised learning task: given a source node *s* and positive and negative training examples, learn a function that assigns edge weights (*i.e.*, random walk transition probabilities) so that a weighted random walk has a higher probability of visiting positive examples than negative examples.

Path based Methods

- Each edge (u, v) has a corresponding feature vector φ_{uv} that describes
  - The nodes u and v (e.g., age, gender, hometown), and
  - The interaction attributes (e.g., when the edge was created, how many messages u and v exchanged, or how many photos they appeared together in).
- Function f_w takes the edge feature vector φ_{uv} as input and computes the corresponding edge weight a_{uv}.
- Goal is to learn parameter w of function f_w .

Path based Methods

- Objective function

$$\min_{w} F(w) = ||w||^2 + \lambda \sum_{d \in D, l \in L} h(p_l - p_d)$$

  - D is the set of positive nodes, L the set of negative nodes,
  - P_i is the probability of visiting node i, and
  - h is a loss function with

$$h(.) = 0 \text{ for } p_d > p_l \text{ and } h(.) > 0 \text{ for } p_d < p_l.$$

Path based Methods

SFU

- Experimental results on Co-authorship dataset

Learning Method	AUC	Prec@20
Random Walk with Restart	0.63831	3.41
Adamic-Adar	0.60570	3.13
Common Friends	0.59370	3.11
Degree	0.56522	3.05
DT: Node features	0.60961	3.54
DT: Network features	0.59302	3.69
DT: Node+Network	0.63711	3.95
DT: Path features	0.56213	1.72
DT: All features	0.61820	3.77
LR: Node features	0.64754	3.19
LR: Network features	0.58732	3.27
LR: Node+Network	0.64644	3.81
LR: Path features	0.67237	2.78
LR: All features	0.67426	3.82
SRW: one edge type	0.69996	4.24
SRW: multiple edge types	**0.71238**	**4.25**

-DT: decision tree
-LR: logistic regression
-SRW: supervised rand.walk

- Most methods have similar performance
- LR is strongest competitor
- SRW outperforms LR

Path based Methods

- Experimental results on Facebook dataset

Learning Method	AUC	Prec@20
Random Walk with Restart	0.81725	6.80
Adamic-Adar	0.81586	7.35
Common Friends	0.80054	7.35
Degree	0.58535	3.25
DT: Node features	0.59248	2.38
DT: Network features	0.76979	5.38
DT: Node+Network	0.76217	5.86
DT: Path features	0.62836	2.46
DT: All features	0.72986	5.34
LR: Node features	0.54134	1.38
LR: Network features	0.80560	7.56
LR: Node+Network	0.80280	7.56
LR: Path features	0.51418	0.74
LR: All features	0.81681	7.52
SRW: one edge type	**0.82502**	6.87
SRW: multiple edge types	**0.82799**	**7.57**

-DT: decision tree
-LR: logistic regression
-SRW: supervised rand.walk

-Unsupervised methods
and LR perform very well
- SRW performs best

Model based Methods

- MF based models [Rennie & Srebo 2005]
  - Social network as a binary matrix.
  - Similar to MF methods for rating prediction.
  - Factorize the network matrix into product of lower rank matrices (representing user factors).
  - Advanced version in [Yang et al. 2011]: factorize user-user matrix and user-item matrix simultaneously.

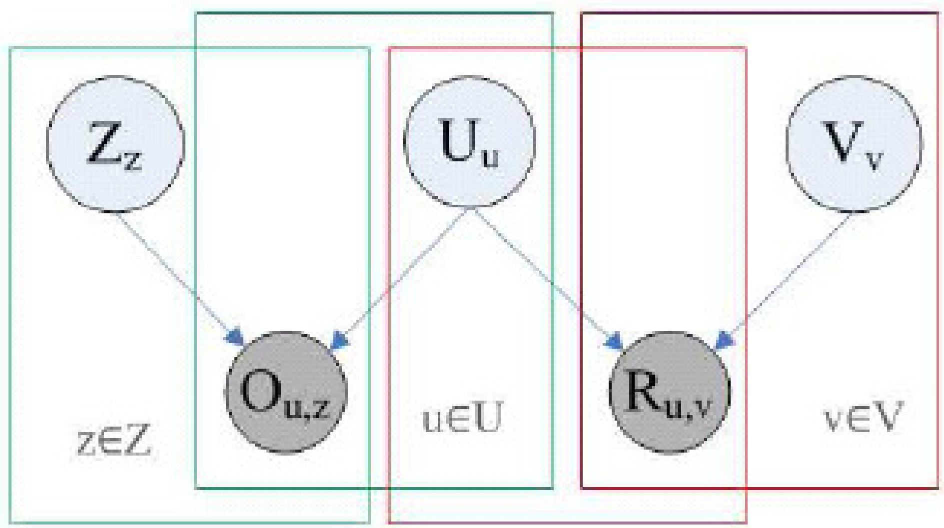
Future Research Directions

- Recommendation in heterogeneous networks with more than 2 entity types

- How to recommend entities of one type to entities of another type?

Future Research Directions

- Collective Matrix Factorization [Singh & Gordon 2008]
- Simultaneously factorize the matrices of all binary relationships.

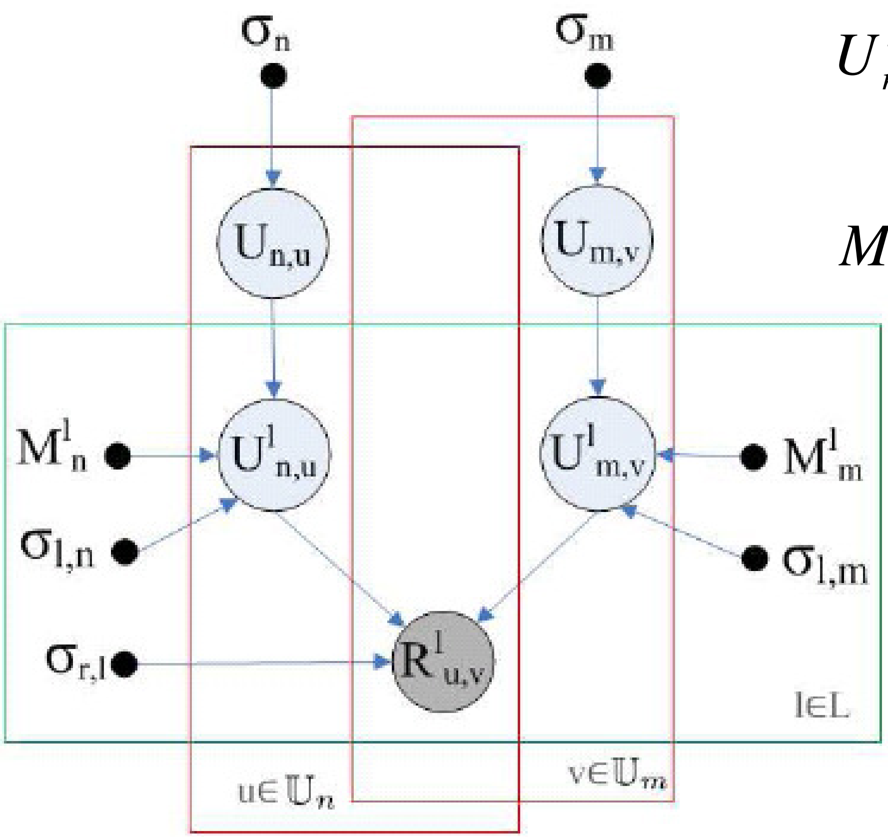


- An entity uses the same latent factors in all of its relationships (contexts).

Future Research Directions

- HeteroMF: Context Dependant Factor Models [Jamali & Lakshmanan 2013]
- Each entity has a base latent factor.
- For each relationship (context) in which its entity type participates, the entity has context specific latent factors.
- Context specific factors are derived from the base factors employing a transfer matrix.

Future Research Directions



U_n : base latent factors

U_n^l : context specific latent factors for context l

M_n^l : transfer matrix for context l

Future Research Directions

- Explanation of social recommendations
- [Papadimitriou et al. 2012] distinguishes the following types of explanations (and their hybrids).
- Human style of explanation, based on similar users.
- Item style of explanation, based on choices made by the user on similar items.
- Feature style of explanation, based on features of items that were previously rated by the user.

Future Research Directions

- Do social explanations work? [Sharma & Cosley 2013]
- Distinguish the following types of explanations:
  - Overall Popularity: The number of Likes by all Facebook users for an artist.
  - Friend Popularity: The number of friends of a user who Like an artist.
  - Random Friend: The name of a random friend, chosen from those that Like an artist.
  - Good Friend: The name of a close friend, chosen from those that Like an artist.
  - Good Friend & Count: A combination of Good Friend and Friend Popularity.

Future Research Directions

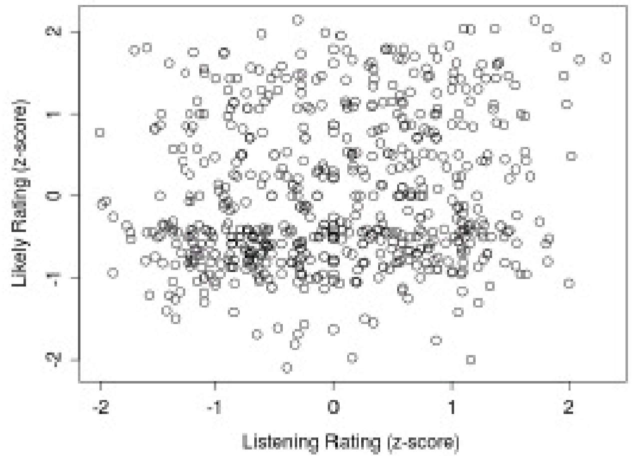
- Persuasiveness of explanations
- For each recommendation, we ask the user how likely (on scale [0..10]) is she to check out the recommended artist.

Explanation Strategy	N	Mean	Std. Dev.
FriendPop	1203	2.12	2.42
RandFriend	1225	2.08	2.49
OverallPop	1191	2.36	2.69
GoodFriend	434	2.52	2.69
GoodFrCount	405	2.71	2.90

- Showing the right friend matters.
- Popularity only matters if user identifies with the crowd.

Future Research Directions

- Informativeness of explanations
- How effective are explanations in directing users to items that receive high consumption ratings?



- The persuasiveness and informativeness of an explanation are quite different.

Future Research Directions

- Privacy of recommendation in social networks
  - How to preserve privacy when employing social networks?
- Recommendation in mobile social networks
  - Can we develop distributed algorithms?
  - How to exploit the user location?
- Recommendation in social networks with user-generated content
  - How can we integrate topic models?

SFU

Thank You!

References

- L. A. Adamic, Eytan Adar. Friends and neighbors on the web. Social Networks, 25(3):211–230, July 2003.
- E. M. Airoldi, D. M. Blei, S. E. Fienberg, E. P. Xing. Mixed membership stochastic blockmodels. Journal of Machine Learning Research, 9:1981–2014, 2008.
- A. Anagnostopoulus, Kumar R., Mahdian M.: Influence and correlation in social networks, KDD 2008.
- L. Backstrom, J. Leskovec. Supervised Random Walks: Predicting and Recommending Links in Social Networks, WSDM 2011.
- J. Golbeck. Computing and Applying Trust in Web-based Social Networks. PhD thesis, University of Maryland College Park, 2005.
- M. Gomez-Rodriguez, J. Leskovec, A. Krause: Inferring networks of diffusion and influence, KDD 2010.
- M. Jamali, M. Ester: A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks, Recsys 2010.
- M. Jamali, M. Ester: TrustWalker : A Random Walk Model for Combining Trust-based and Item-based Recommendation, KDD 2009.
- M. Jamali, T. Huang, M. Ester: A Generalized Stochastic Block Model for Recommendation in Social Rating Networks, ACM RecSys 2011.
- M. Jamali, L. Lakshmanan: HeteroMF: Recommendation in Heterogeneous Information Networks using Context Dependent Factor Models, WWW 2013.

References

- G. Jeh, J. Widom: Simrank: a measure of structural-context similarity, KDD 2002.
- L. Katz: A new status index derived from sociometric analysis, Psychometrika, 18(1):39–43, 1953.
- Y. Koren, R. Bell R., C. Volinsky: Matrix factorization techniques for recommender systems, IEEE Computer, 2009.
- D. Liben-Nowell, J. Kleinberg. The Link Prediction Problem for Social Networks, CIKM 2003.
- D. Lin: An information-theoretic definition of similarity, ICML 1998.
- H. Ma, I. King, M. R. Lyu: Learning to recommend with social trustensemble, SIGIR 2009.
- H. Ma, M. R. Lyu, Irwin King: Learning to recommend with trust and distrust relationships, RecSys 2009.
- H. Ma, H. Yang, M. R. Lyu, I. King. Sorec: social recommendation using probabilistic matrix factorization, CIKM 2008.
- P. Massa, P. Avesani. Trust-aware recommender systems, RecSys 2007.
- S.A. Myers, J. Leskovec: On the convexity of latent social network inference, NIPS 2010.
- P.R. Monge, N.S. Contractor: Theories of communication networks, Oxford University Press, 2003.
- M. E. J. Newman. Clustering and preferential attachment in growing networks, 2001.
- J.-Y. Pan, H.-J. Yang, C. Faloutsos, P. Duygulu. Automatic multimedia cross-modal correlation discovery, KDD 2004.

References

- A. Papadimitriou, P. Symeonidis, Y. Manolopoulos: A generalized taxonomy of explanations styles for traditional and social recommender systems, Data Mining and Knowledge Discovery, 24(3), 555-583, 2012.
- C. Posse: Key Lessons Learned Building Recommender Systems for Large-Scale Social Networks, KDD 2012.
- J.D.M. Rennie, N. Srebro. Fast maximum margin matrix factorization for collaborative prediction, ICML 2005.
- A. P. Singh, G. J. Gordon. Relational learning via collective matrix factorization, KDD 2008.
- A. Sharma, D. Cosley: Do social explanations work? Studying and modeling the effects of social explanations in recommender systems, WWW 2013.
- H. Steck: Training and testing of recommender systems on data missing not at random, KDD 2010.
- S. Wasserman, K. Faust: Social network analysis: methods and applications, Cambridge University Press, 1994 .
- S.H. Yang, B. Long, A. Smola, N. Sadagopan, Z. Zheng, H. Zha: Like like alike - Joint Friendship and Interest Propagation in Social Networks, WWW 2011.
- X. Yang, H.Steck, Y. Guo, Y. Liu: On top-k recommendation using social networks, RecSys 2012.