Hete-CF: Social-Based Collaborative Filtering Recommendation using Heterogeneous Relations

Chen Luo[1], Wei Pang[2], Zhe Wang[1], Chenghua Lin[2]

[1]Jilin University, China

[2]Aberdeen University, UK

http://www.chen-luo.com/

Outline

- Background and Motivation
- The Hete-CF Model
- Experimental Result
- Conclusion

CF-based Recommendation

- Use user-item rating matrix $\mathbb{R}$ to make recommendation.

$$\min_{U,V} \sum_{i=0}^{m} \sum_{j=0}^{n} (U_i^T V_j - \mathbb{R}_{i,j})^2 + \lambda(\|U\|_F^2 + \|V\|_F^2)$$

- *U* and *V* are lower rank representations of users and items.
- Then the ratings can be calculated as

$$\hat{R_{ij}} = U_i^T V_j$$

- **Suffered by Data Sparsity and Cold start problems.**

Dec 14-17, 2014 ICDM SHENZHEN, CHINA

Social Recommendation

- Using the information from social networks to help recommendation.
  - Trust-MF (IJCAI 2013) Uses trust relationship **between users** to improve recommendation
  - Hete-MF (IJCAI-HINA 2013) Uses **item (Entity) similarity** information to improve recommendation
- There are also another types of relations which can improve the recommendation performance.
  - **Other Relations between users and items.**

A Real Example

- **Recommendation Task:** Recommend Conference for Author.
- **Target Recommendation Relation**:

  Author-Paper-Conference
- **Social Information for improvement:**
  - Relations between users: Author-Paper-Author.
  - Relations between items: Conference-Topic-Conference.
  - Relations between users and items: Author-Topic-Author-Paper-Conf.
- **We want to design a model to use all these three types of relation**

Outline

- Background and Motivation
- The Hete-CF Model and Algorithm
- Experimental Result
- Conclusion

Model Overview

- Part 1: Modeling Relations between users
  - A good recommendation result should consider the relations between users.
- Part 2: Modeling Relations between items
  - A good recommendation result should consider the relations between items
- Part3: Modeling Relations between users and items
  - Some other relations between users and items can improve the recommendation result.

A Unified Model.

Part 1: Modeling Relations between users

- Using Graph Regularization:

$$\min_{U,A} \sum_{k=0}^{N_A} \alpha_k \sum_{i=0}^{n} \sum_{j=0}^{n} S_A^k(i,j) \, \|U_i - U_j\|_F^2$$

- ***A:*** weight for each relations between users introduced.
- S_A^k is the similarity matrix between users, can be calculated as follow:

$$S_A^k(i,j) = S_{P_k^A}^{PS}(i,j)$$

- The P_k^A can be calculated using ***PathSim*** *(VLDB'11)*

Dec 14-17, 2014 ICDM SHENZHEN, CHINA

Part 2: Modeling Relations between items

- Using Graph Regularization:

$$\min_{V,B} \sum_{k=0}^{N_B} \beta_k \sum_{i=0}^{m} \sum_{j=0}^{m} S_B^k(i,j) \left\| V_i - V_j \right\|_F^2$$

- ***B:*** weight for each relations between items introduced.
- S_B^k is the similarity matrix between items, can be calculated as follow:

$$S_B^k(i,j) = S_{P_k^B}^{PS}(i,j)$$

- The $S_{P_k^A}^{PS}(i,j)$ can be calculated using ***PathSim***

Part 3: Modeling Relations between users and items

- Using Collaborative Model:

$$\min_{U,V,W} \sum_{k=0}^{N_W} w_k \sum_{i=0}^{n} \sum_{j=0}^{m} (U_i^T V_j - \mathbb{R}_{i,j}^k)^2$$

- ***W:*** weight for each relations between users and items introduced.
- $\mathbb{R}^k$ is the relation graph of the *k-th* relation between users and items:

$$\mathbb{R}^k(i,j) = S_{P_k^W}^{PS}(i,j)$$

- The $S_{P_k^W}^{PS}(i,j)$ can be calculated using ***PathSim***

Dec 14-17, 2014 ICDM SHENZHEN, CHINA

JILIN UNIVERSITY-CHINA 1946

The Unified Overview

- Model

$$\min_{U,V,A,B,W} \sum_{i=0}^{m}\sum_{j=0}^{n}(U_i^T V_j - \mathbb{R}_{i,j})^2$$

$$+\sum_{k=0}^{N_A}\alpha_k \sum_{i=0}^{n}\sum_{j=0}^{n} S_A^k(i,j)\,\|U_i - U_j\|_F^2$$

Modeling the relations between users

$$+\sum_{k=0}^{N_B}\beta_k \sum_{i=0}^{m}\sum_{j=0}^{m} S_B^k(i,j)\,\|V_i - V_j\|_F^2$$

Modeling the relations between items

$$+\mu \sum_{k=0}^{N_W} w_k \sum_{i=0}^{n}\sum_{j=0}^{m}(U_i^T V_j - \mathbb{R}_{i,j}^k)^2$$

Modeling the relations between users and items

$$+\lambda(\|U\|_F^2 + \|V\|_F^2 + \|A\|_F^2 + \|B\|_F^2 + \|W\|_F^2)$$

Dec 14-17, 2014 ICDM SHENZHEN, CHINA
JILIN UNIVERSITY CHINA 1946

The Learning Algorithm

- A **Two-Step Iteration Method**.
  - The predicted rating vector ***U, V*** *and the weight for each relation* ***A,B,W*** mutually enhance each other.
- Optimize ***U,V*** Given ***A,B,W***
  - Becomes to a traditional CF model
  - Using SGD to obtain *U,V*
- Optimize ***A,B,W*** Given ***U,V***
  - Becomes a linear model
  - Also using SGD to obtain A,B,W

Input: A heterogeneous information network $G = \langle O, E, W \rangle$. Three sets of meta-paths between user and item, users, and items. The user-item rating matrix $\mathbb{R}$. Parameter λ, α^s.

Output: The rating Matrix $\hat{R}$;

Initialize U, V, A, B, W randomly;

while *not reaching the inner* U, V, A, B, C *difference threshold* **do**

    while *not reaching the inner* U, V *difference threshold* **do**

        Update U, V using Eqs.(16) and (17);

    end

    while *not reaching the inner* A, B, C *difference threshold* **do**

        Update A, B, W using Eqs. (22), (23) and (24);

    end

end

The prediction rating is $\hat{R_{ij}} = f(U_i^T V_j)$.;

return $\hat{R_{ij}}$;

Algorithm 1: Hete-CF

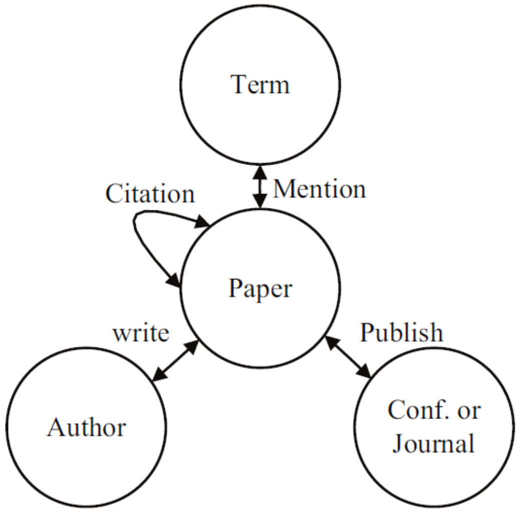
Dec 14-17, 2014 ICDM SHENZHEN, CHINA

Outline

- Background and Motivation
- The Hete-CF Model and Algorithm
- Experimental Result
- Conclusion

Experimental Setup

- **Datasets**
  - Sub-DBLP Dataset
  - 261 Journal and 313 conference paper
- **Task**
  - Recommend ***Conference*** to ***Author***
- **Cross Validation**
  - 5-fold cross validation
  - 40%(60%) as training
  - 60%(40%) as testing



- **Baseline**
  - UserMean, ItemMean, NMF
- **State-of-the-art**
  - Trust-MF (IJCAI-13),
  - Hete-MF (IJCAI-HINA 13)
- **Evaluation**
  - MAE
  - RMSE

Experimental Result

- Result

TABLE I: Algorithm Performance Comparison in DBLP

% Training	Feature	Evaluation	UserMean	ItemMean	NMF	Trust-MF	Hete-MF	Hete-CF
40%	$d=5$	MAE	0.942 ± 0.02	1.065 ± 0.02	2.156 ± 0.02	0.831 ± 0.01	0.931 ± 0.02	$\mathbf{0.831 \pm 0.02}$
		RMSE	1.216 ± 0.01	1.123 ± 0.02	2.394 ± 0.01	1.013 ± 0.02	1.105 ± 0.01	$\mathbf{1.002 \pm 0.03}$
	$d=10$	MAE	0.943 ± 0.03	0.948 ± 0.01	2.194 ± 0.03	0.887 ± 0.01	0.901 ± 0.01	$\mathbf{0.859 \pm 0.01}$
		RMSE	1.138 ± 0.02	1.256 ± 0.04	2.292 ± 0.02	1.083 ± 0.03	1.114 ± 0.03	$\mathbf{1.056 \pm 0.02}$
60%	$d=5$	MAE	0.948 ± 0.02	0.919 ± 0.01	2.131 ± 0.04	$\mathbf{0.812 \pm 0.02}$	0.891 ± 0.02	0.831 ± 0.02
		RMSE	1.132 ± 0.02	1.157 ± 0.01	2.385 ± 0.01	$\mathbf{0.907 \pm 0.02}$	1.010 ± 0.03	0.938 ± 0.02
	$d=10$	MAE	0.932 ± 0.03	0.978 ± 0.03	2.184 ± 0.02	0.873 ± 0.03	0.881 ± 0.01	$\mathbf{0.856 \pm 0.02}$
		RMSE	1.154 ± 0.02	1.143 ± 0.02	2.275 ± 0.01	1.051 ± 0.01	1.013 ± 0.02	$\mathbf{0.994 \pm 0.03}$

Outline

- Background and Motivation
- The Hete-CF Model and Algorithm
- Experimental Result
- Conclusion

Conclusion

- We proposed Hete-CF a CF-based social recommendation method using heterogeneous relations
- Hete-CF utilize three types of relations in a heterogeneous network.
- The experimental result showed the effectiveness of Hete-CF

Chen Luo
Master Candidate from Jilin University
Home Page: http://www.chen-luo.com/

Q&A