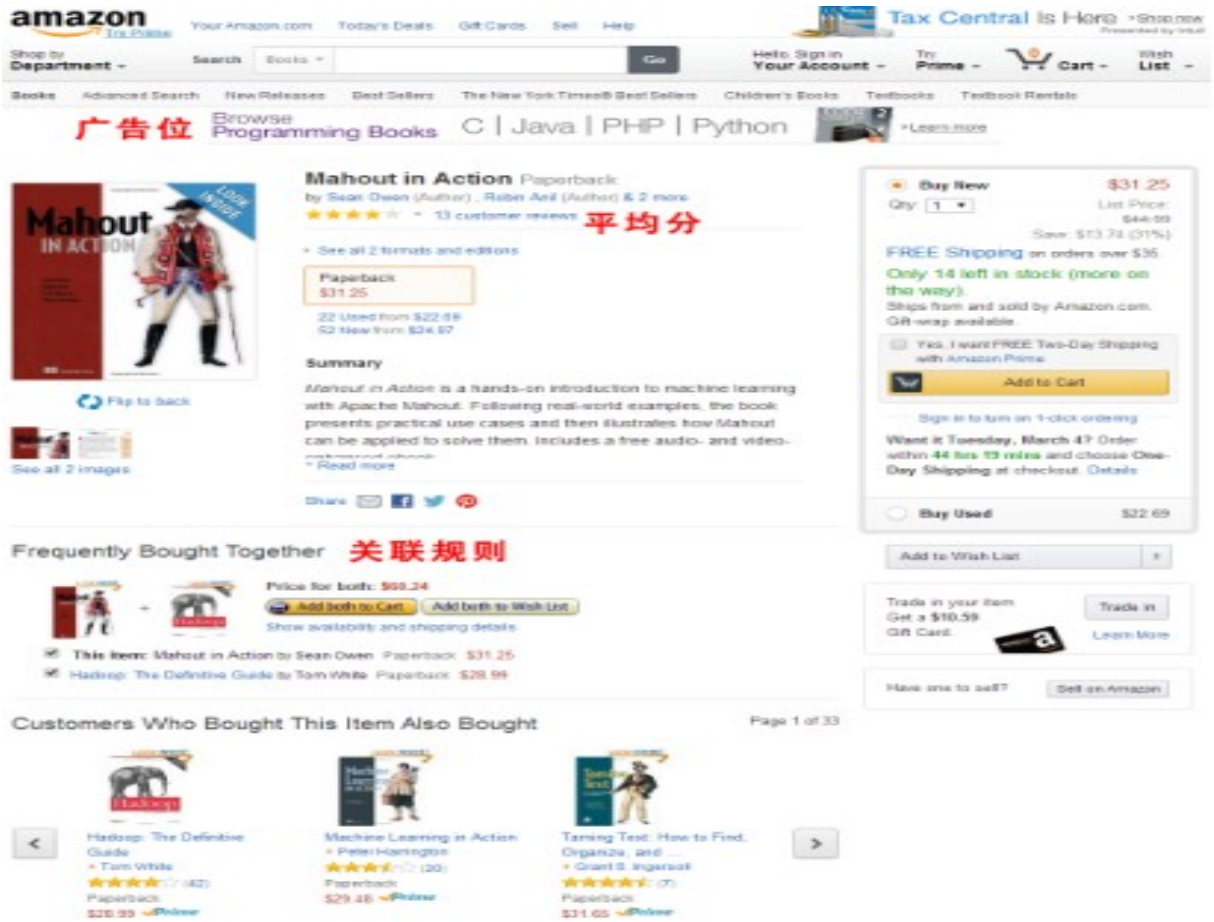
Item-Based and User-Based Collaborative Filtering

Tao Shan

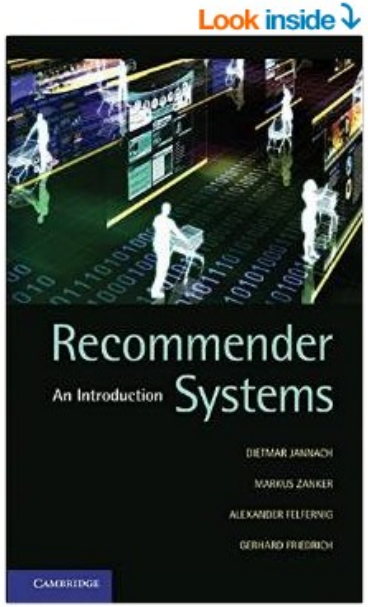
1.Recommend System

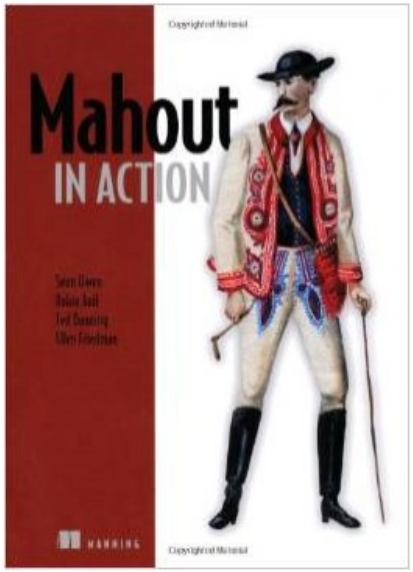
- Amazon.
- Ebay.



2.How the System know it?

- Material for learning:





- http://portal.acm.org/citation.cfm?id=1070751
- http://portal.acm.org/citation.cfm?id=372071

3.Collaborative Filtering

- Basic Idea:In huge amount of user find some people share similar taste with you.Recommend the things they like to you.

- Problem:
- How to find out these people?
- How to sort the item ?

3.Collaborative Filtering

- Implementation:

- First Step:
  - Collect user's preference.
- Second Step:
  - Find out similar item or user.
- Third Step:
  - Recommed.

4.First Step: Collect Data.

- Build data matrix:
- Item-Based:

	User-1	User-2	User-3	User-4	User-5
Item-1	1	1	0	0	1
Item-2	0	1	1	0	1

- User-Based:

	Item-1	Item-2	Item-3	Item-4	Item-5
User-1	1	1	0	1	0
User-2	1	1	1	1	0

5.Item-CF and User-CF

- Item-CF:
  - Focus on the item.Recommend the item which are similar with each other to you.
- User-CF:
  - According to the user whose taste is similar to you.Recommed the item he likes to you.

6.Find out the similarity item or user.

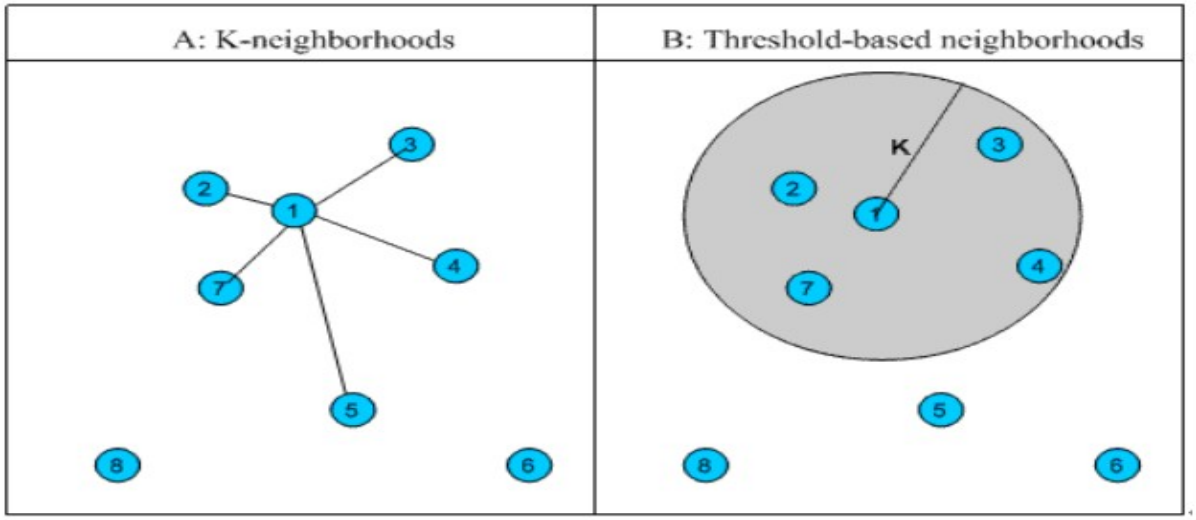
- According to the matrix we can caldulate the similarity between every item or user.

- We call it "Co-occurrence degree"

7.Recommend

- So many cadidate items or users.Only chose some of them ot recommend.

- Threshold-based neighborhoods or K-neighborhoods



8.Example

Question:

All the e-business companies "Amazone ,Ebay , Alibaba" choose the "Item-CF" as their Recommend System algorithem. Why?

Our example:

Movie recommend system.

8.Example

- Origin Data:
  - Fome Mahout in Action Ch.6
- 3 field
  - User-ID
  - Movie-ID
  - Score

```
1,101,5.0
1,102,3.0
1,103,2.5
2,101,2.0
2,102,2.5
2,103,5.0
2,104,2.0
3,101,2.0
3,104,4.0
3,105,4.5
3,107,5.0
4,101,5.0
4,103,3.0
4,104,4.5
4,106,4.0
5,101,4.0
5,102,3.0
5,103,2.0
5,104,4.0
5,105,3.5
5,106,4.0
```

8.Example

- 1.Build Score Matrix:

	101	102	103	104	105	106	107
1	5	3	2	0	0	0	0
2	2	2.5	5	2	0	0	0
3	2	0	0	4	4.5	0	5
4	5	0	3	4.5	0	4	0
5	4	3	2	4	3.5	4	0

8.Example

- 2.Build Co-occurrence

	[101]	[102]	[103]	[104]	[105]	[106]	[107]
[101]	5	3	4	4	2	2	1
[102]	3	3	3	2	1	1	0
[103]	4	3	4	3	1	2	0
[104]	4	2	3	4	2	2	1
[105]	2	1	1	2	2	1	1
[106]	2	1	2	2	1	2	0
[107]	1	0	0	1	1	0	1

8.Example

- Matrix Mutiply the Score Vector:

	101	102	103	104	105	106	107		U3		R
101	5	3	4	4	2	2	1		2.0		40.0
102	3	3	3	2	1	1	0		0.0		18.5
103	**4**	**3**	**4**	**3**	**1**	**2**	**0**	**X**	**0.0**	**=**	**24.5**
104	4	2	3	4	2	2	1		4.0		40.0
105	2	1	1	2	2	1	1		4.5		26.0
106	2	1	2	2	1	2	0		0.0		16.5
107	1	0	0	1	1	0	1		5.0		15.5

9.Evaluation

	Like	Not Like
Recommend	A	C
Not Recommend	B	D

PRECISION: A/A+C
RECALL:A/A+B

Find a balance between two value.

Thank You