

Group 7

Privacy and Security Issues in Recommender Systems

Bing HU 1155151103

Heng LENG 1155152543

Index

C O N T E N T S

01

Introduction

04

Security Issues in RS

02

Recommendation
Algorithms

05

Future work

03

Privacy Issues in RS



0 1

Introduction

Recommender system

-- Concept



Potential solution to the problem of
information overload

recommendation system

“

Personalized system
that recommends
the information and
products

Recommender system

-- 3 modules



user modeling module



recommendation object modeling
module



recommendation algorithm module



0 2

Recommendation Algorithms

Recommendation Algorithm

-- Types



CBR (content based algorithm)



CF (collaborative filtering)



knowledge-based &
hybrid algorithm



Recommendation algorithms is the most important part of the whole recommender system, and largely decides the type of recommender system and the performance.

Recommendation Algorithm -- Work flow



building user profiles



computing between users



find the top n

0 3

Privacy Issues in RS

Concepts

recommendation server

- **semi-trusted**
- **malicious**

dilemma:

- **improve the accuracy and usability**
- **Or privacy exposure ?**

Privacy Preserving Techniques

cryptographic methods



distributed storage

More details in the paper

0 4

Security Issues in RS

Concept of shilling attack

	Item1	Item2	Item3	Item4	Item5	Item6	Correlation with Alice
Alice	5	2	3	3		?	
User1	2		4		4	1	-1.00
User2	3	1	3		1	2	0.76
User3	4	2	3	1		1	0.72
User4	3	3	2	1	3	1	0.21
User5		3		1	2		-1.00
User6	4	3		3	3	2	0.94
User7		5		1	5	1	-1.00
Attack1	5		3		2	5	1.00
Attack2	5	1	4		2	5	0.89
Attack3	5	2	2	2		5	0.93
Correlation with Item6	0.85	-0.55	0.00	0.48	-0.59		

Figure 1: an example of a user-based collaborative recommender system being affected by shilling attack.[8]

Purpose of the attack

- Affecting the recommendation frequency of the target item.
- disrupt the recommendation accuracy

Knowledge needed

- item information, user information, scoring information, and the recommendation algorithm used

Cost of attack

- knowledge cost and execution cost.

Main attack model

01

Random attack

Target item is given the predetermined value and filled items are given random values within the rating scale with a distribution centered around the average value of all user ratings in all items.

02

Average attack

The difference between the average attack and the random attack model in the attack profile is that the rating of each filled item of the former is the average value

03

Bandwagon attack

Its basic idea is based on Zip's law, that is, a few items can attract the attention of most people

04

Segmented attack

The goal of the segmented attack model is to recommend target items for a specific user group

Attack Detection



Supervised learning

Training the detector with a known category of users as a reference is an intuitive idea when thinking about the detection of a shilling attack.



Unsupervised learning

capable of new or confused shilling attacks

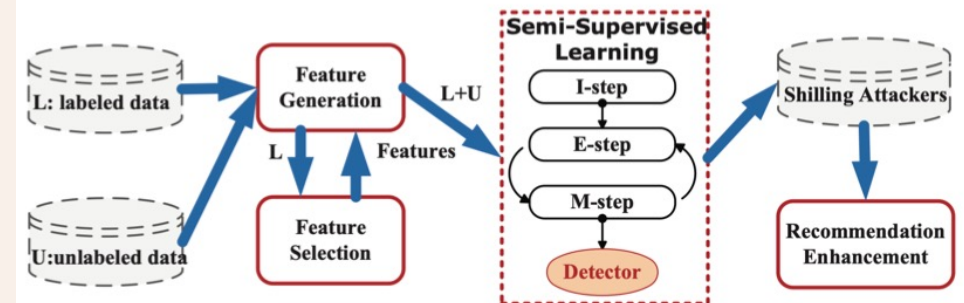


Semi-supervised learning

precious labeled user data & distribution pattern shown unlabeled users

Attack detection

Shilling attack detection is essentially a classification problem. According to the use of prior knowledge, detection algorithms can be divided into three categories: supervised learning, unsupervised learning and semi-supervised learning.



Attack defense



- Increase the cost of attack
- Seek for robust recommendation algorithms

Future Work

Privacy:

- find a lightweight and verifiable security model
- when some cryptographic methods like public key encryption have to be used to preserve privacy, could also find a good way to reduce the complexity of operation

Future Work

Security:

- Comparative analysis of the defense
- Further improvement of recommendation algorithm
- Development of detection tools for shilling attack
- What kind of role users can play
- Protection of new items

Group 7

Thank you for watching!

Bing HU 1155151103

Heng LENG 1155152543