

# アイテム依存の 撮動化プライバシー保護情報推薦

† 望月 安菜

† 菊池 浩明


† 東海大学

# 情報推薦



## 耳をすませば [DVD]

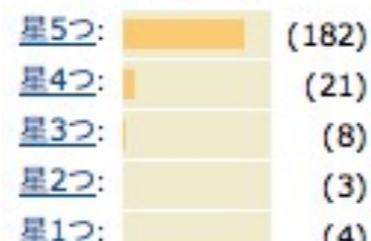
近藤喜文 (監督) | 形式: DVD

★★★★☆ (218件のカスタマーレビュー)  いいね (22)

### カスタマーレビュー

#### 耳をすませば [DVD]

218レビュー



おすすめ度

★★★★☆ (218件のカスタマーレビュー)

あなたの意見や感想を教えてください

[自分のレビューを作成する](#)

## これにも注目

最近チェックした商品

同じテーマの商品



猫の恩返し/ギブリーズ  
episode 2

森田宏幸, 池脇千鶴, 袴田吉彦, ...

VHS



猫の恩返し/ギブリーズ  
episode2

森田宏幸, 百瀬義行, 宮崎駿

DVD

¥4,935 ¥ 3,663



崖の上のポニョ [DVD]

宮崎駿

DVD

¥4,935 ¥ 3,150



ゲド戦記 [DVD]

宮崎吾朗

DVD

¥4,935 ¥ 2,980



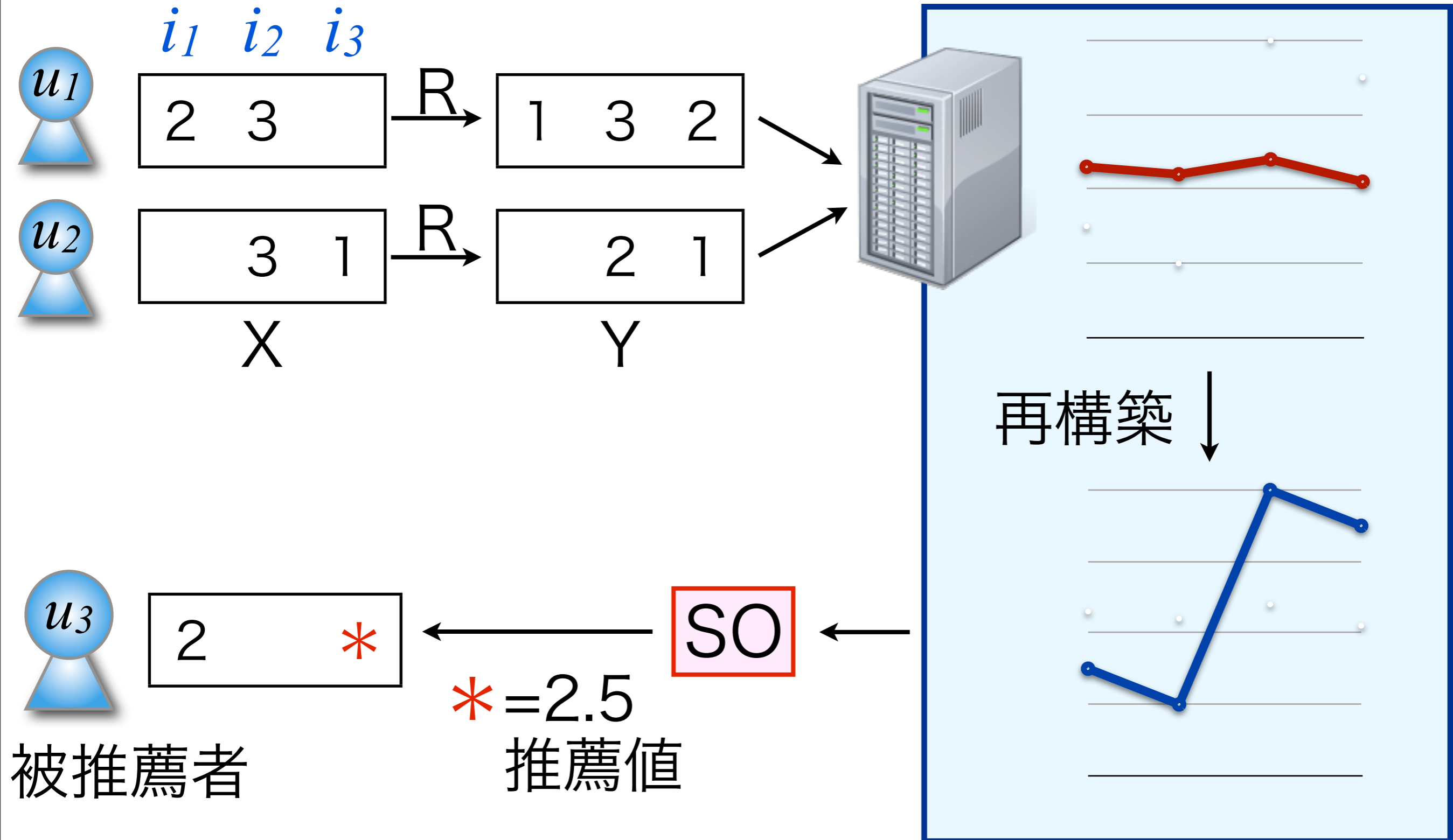
ハウルの動く城 [DVD]

宮崎駿

DVD

¥4,935 ¥ 3,700

# 摂動化プライバシー保護データマイニング



# 摂動化の問題点 (アイテム不変摂動化)

## 猫の恩返し



攪乱不足  
(too small)

	$i_1$	$i_2$	$i_3$	$i_4$
$u_1$	5	5		1
$u_2$	5	4		
$u_3$	5	3		
$u_4$	5		2	5

摂動化



維持確率  $p$

	$i_1$	$i_2$	$i_3$	$i_4$
$u_1$	1	5		1
$u_2$	5	2	3	
$u_3$	2	3	2	2
$u_4$				5

攪乱しすぎ  
(too much)



## となりのトトロ

# 2つの評価尺度

MAE

誤差

相互情報量

特定

匿名

正常

変化 ←

$p$  : 維持確率

→

維持

# 提案・研究目的

## ❖ アイテム依存した摂動化情報推薦方式を提案

- 精度の向上
- プライバシー向上

## ❖ アイテム不変

	$i_1$	$i_2$	$i_3$
$u_1$	4	5	
$u_2$	3		
$u_3$	5	4	
$u_4$			1
$u_5$	5	2	
維持確率	0.5		

## ❖ アイテム依存

	$i_1$	$i_2$	$i_3$
$u_1$	4	5	
$u_2$	3		
$u_3$	5	4	
$u_4$			1
$u_5$	5	2	
維持確率	0.8	0.6	0.2

## ❖ S. Zhang et al.

- 全てのアイテムを一様にアイテムを一様にアイテム不変に摂動化することは、精度の面でも、プライバシーの面でも効率が悪いことを指摘
- ユーザとの対話で維持確率を決め、特異値分解 (SVD) を用いて摂動化を行う

## ❖ 菊池, 五十嵐ら

- 分析に重要で変化させたくないセンシティブ属性を保持する
- 通常の攪乱・再構築に比べ計算コストを削減することができる

# 提案アプローチ

I. アイテム毎の電子投票でスパース率 $s$ を求める

II.  $s$  について維持確率 $p(s_i)$ を決める

III.  $p(s_i)$  でアイテム $i$  について再構築 [DICOMO2011]

$$p(s_i) = 1 - s_i$$

IV. 摂動化Slope One 推薦アルゴリズムを実行

[CSS2011]

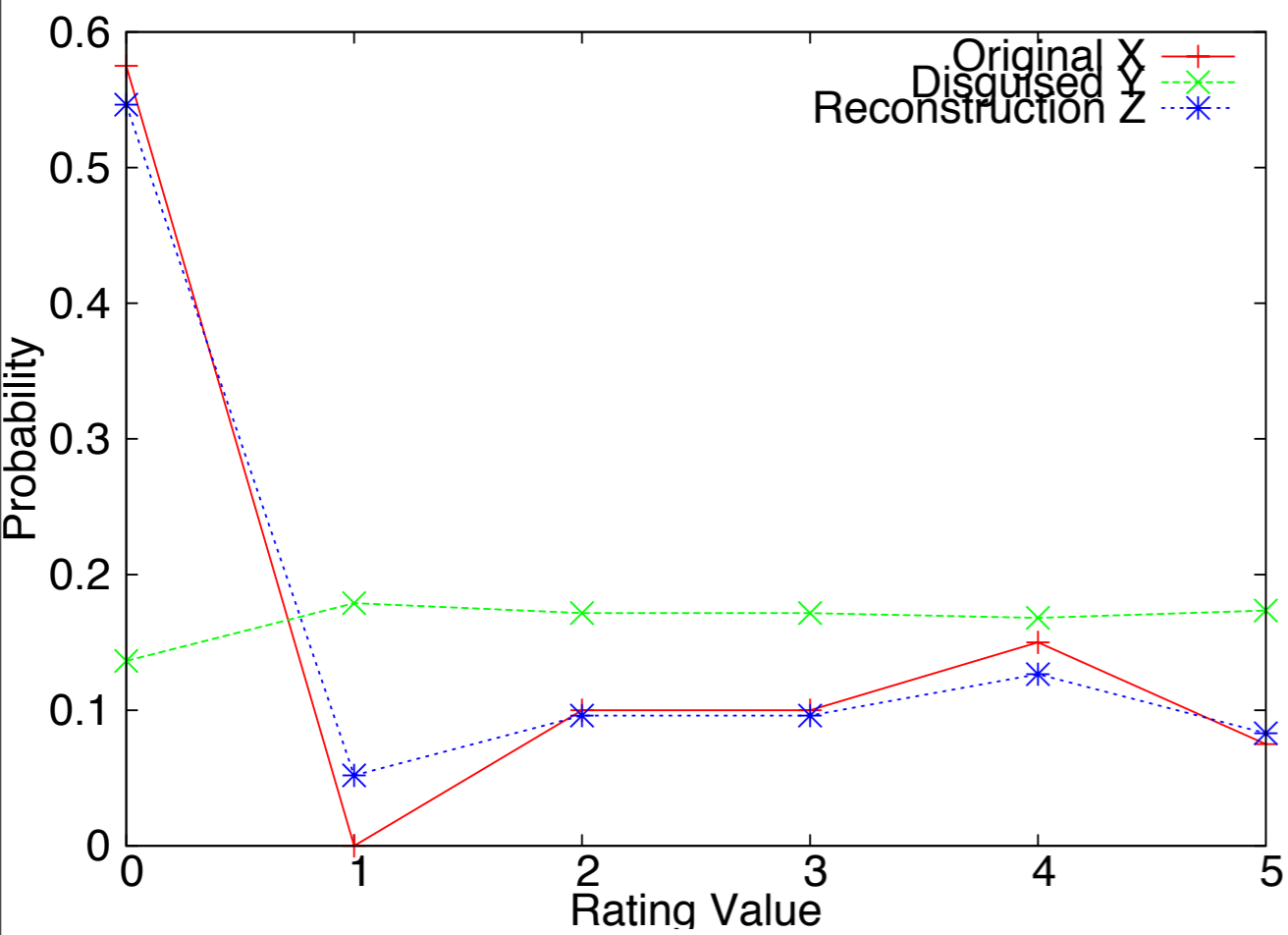


# 維持確率の決め方

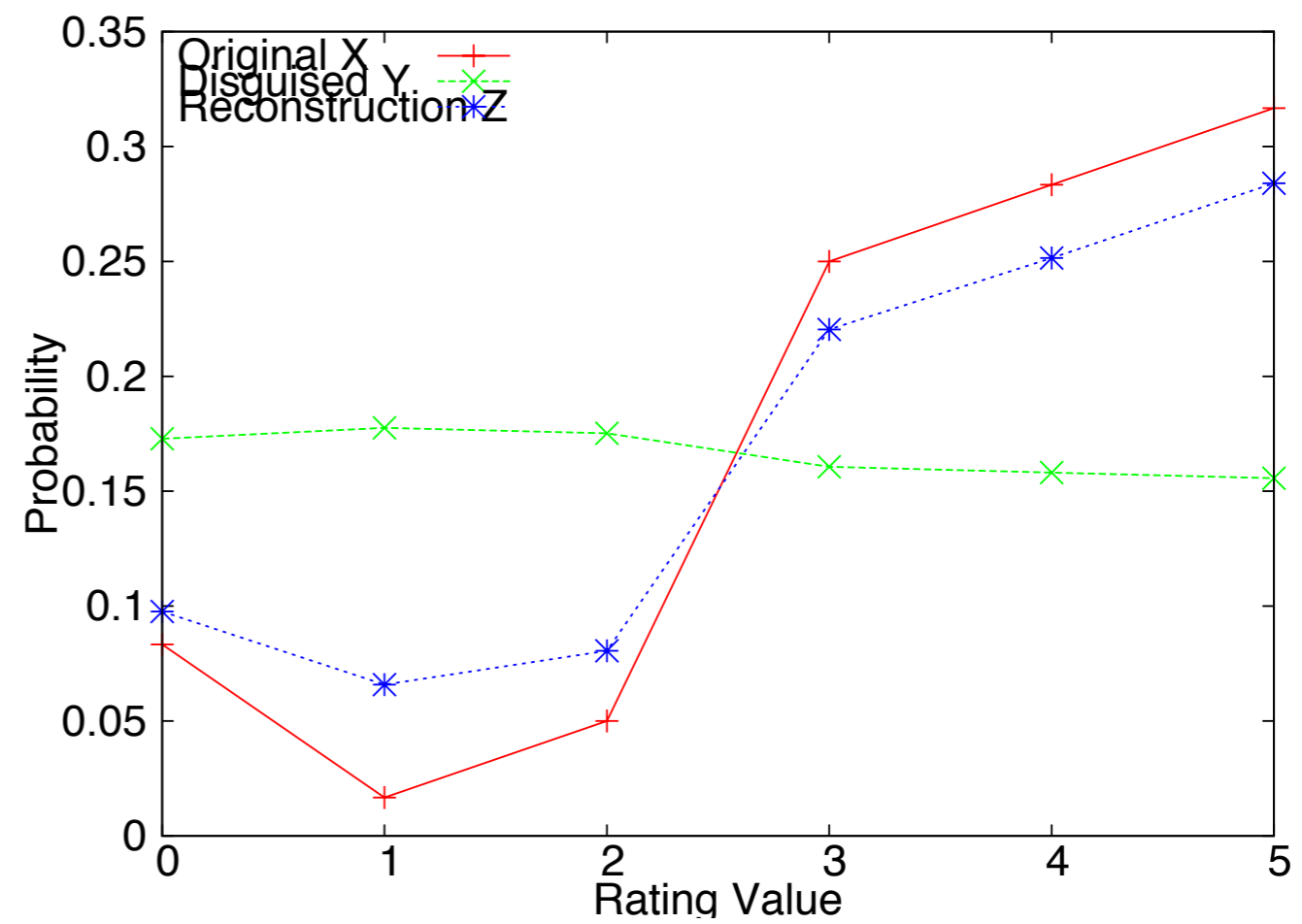
	欠損・少	欠損・多
一様評価	 <p><math>p \rightarrow</math> 大 となりのトトロ</p>	 <p><math>p \rightarrow</math> 小 紅の豚</p>
多様評価	 <p><math>p \rightarrow</math> 小 崖の上のポニョ</p>	 <p><math>p \rightarrow</math> 中 猫の恩返し</p>

# 誤差とプライバシー

$p=0.1$



$p=0.8$

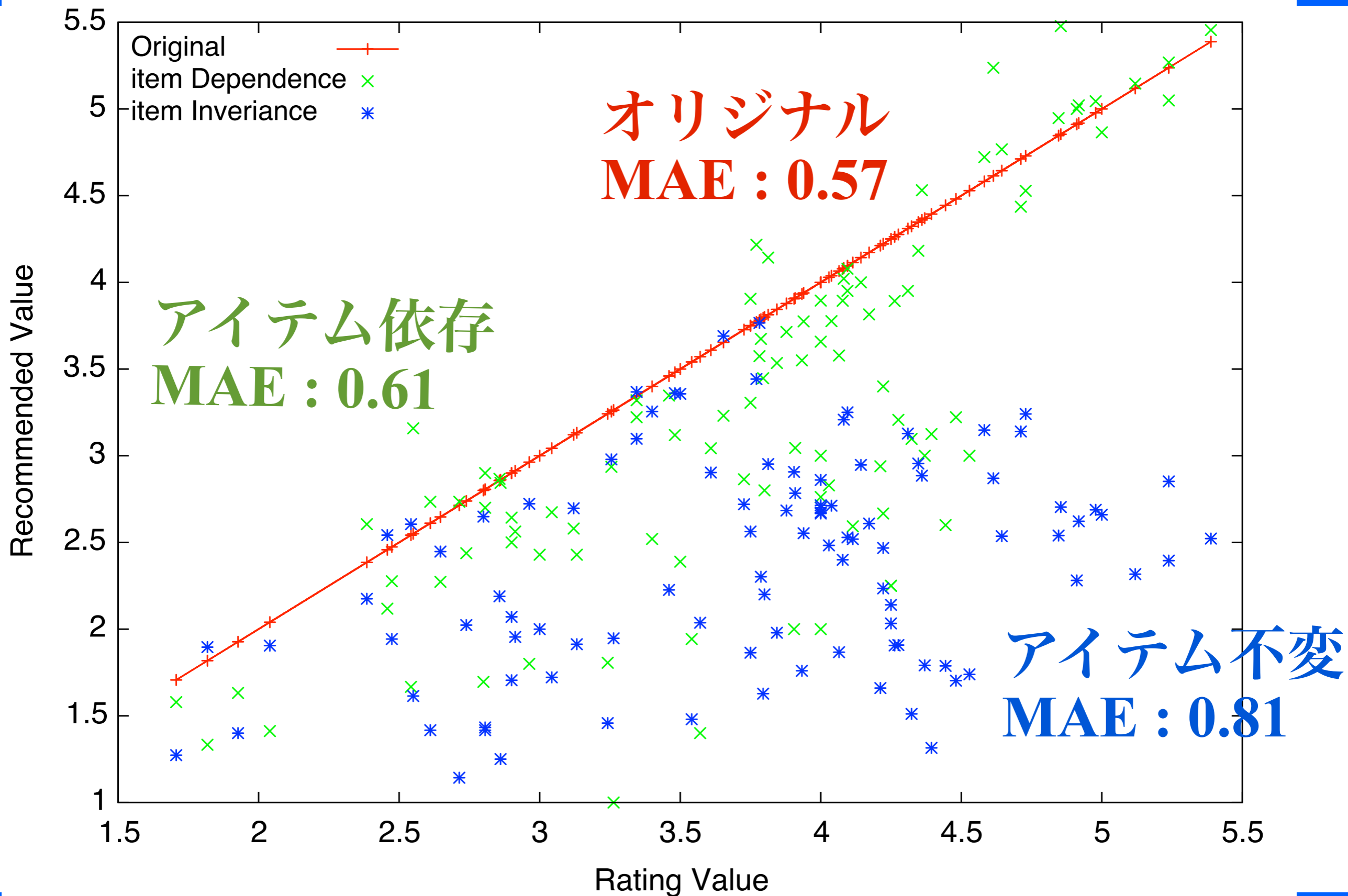


- ❖ アイテム依存推薦の有効性を検証
- ❖ 使用データ：MovieLens data set
  - 評価数 10万件
  - $n=943, m=1628$
  - ランダムに  $n=10, m=10$

# 実験データ

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_7$	$i_8$	$i_9$	$i_{10}$
$u_1$	5	4	5	5		3			3	2
$u_2$	5	3		5	3	4			5	
$u_3$	5	4	5	3	1	4	4		3	
$u_4$	5	5	5	4		3	4			
$u_5$	5	3	3	5	3	4		4		2
$u_6$	5		5	4	3	4			5	
$u_7$	5	4	5	4	3	3				3
$u_8$	5	3	4	3	2	3	3	2	4	
$u_9$	4	5	5	4	3	4				4
$u_{10}$	4	2	4	2		4	2	5	4	
$p$	0.8						0.1			

# 精度



## ❖ まとめ

- アイテム依存の摂動化を提案
- 維持確率を変えることによって、計算コストを削減できる

## ❖ 今後の課題

- 適切な維持確率
- アイテム依存維持確率を使った情報推薦