



## 情報推薦における近傍ユーザの自動選択

|       |  |
|-------|--|
| メタデータ | 言語: jpn<br>出版者: 計測自動制御学会<br>公開日: 2016-02-17<br>キーワード (Ja): 情報推薦, 変数選択法<br>キーワード (En):<br>作成者: 三浦, 功輝, 武内, 充, 岡田, 吉史<br>メールアドレス:<br>所属: |
| URL   | <a href="http://hdl.handle.net/10258/3858">http://hdl.handle.net/10258/3858</a>  |

## 情報推薦における近傍ユーザの自動選択

|     |   |
|-----|---|
| 著者  | 三浦 功輝, 武内 充, 岡田 吉史  |
| 雑誌名 | 計測自動制御学会システム・情報部門学術講演会講演論文集   |
| 巻   | 2015  |
| 発行年 | 2015-11-18  |
| URL | <a href="http://hdl.handle.net/10258/3858">http://hdl.handle.net/10258/3858</a> |

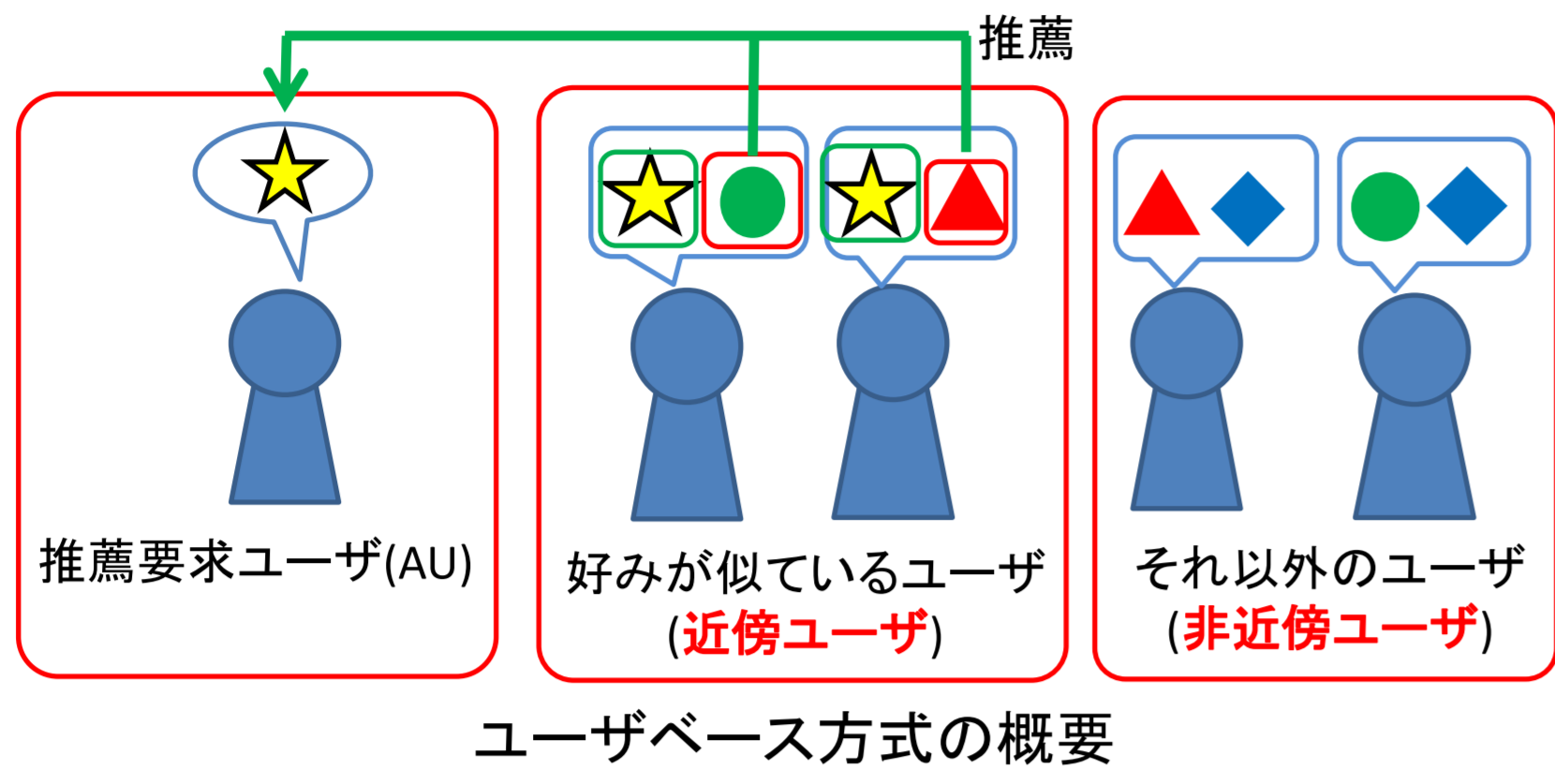


# 情報推薦における近傍ユーザの自動選択

三浦功輝, 武内充, 岡田吉史 (室蘭工業大学)

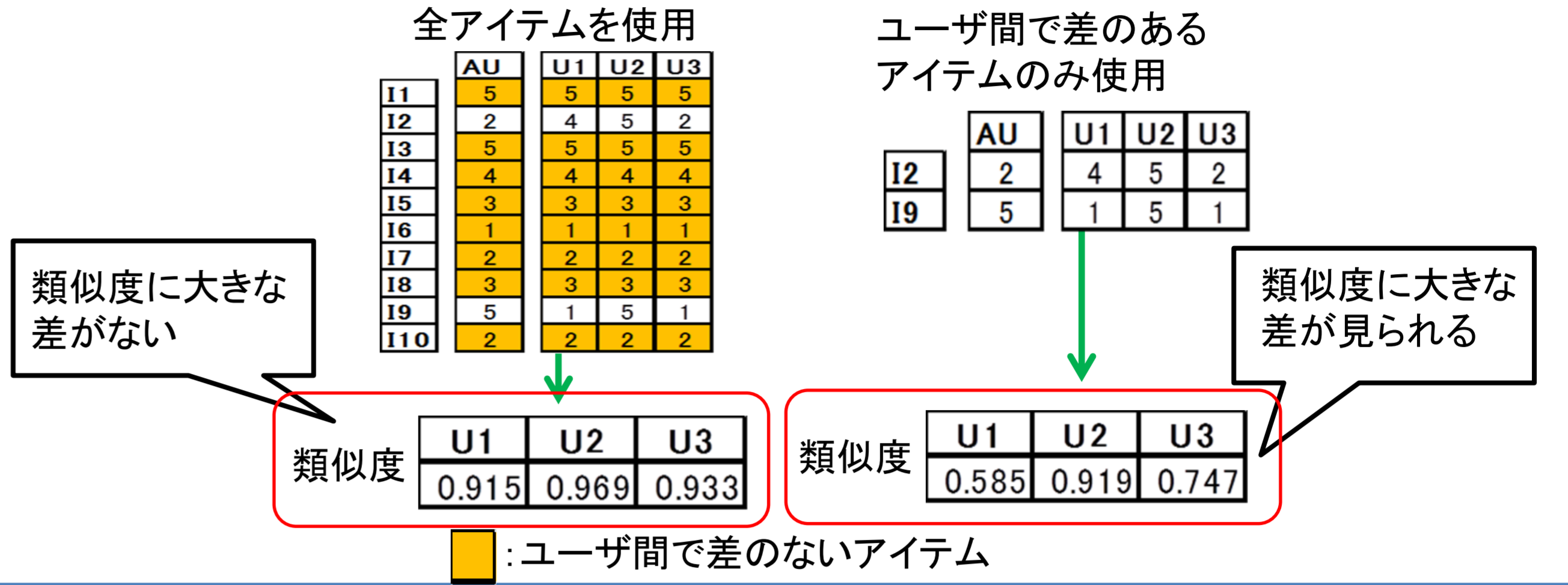
## 研究背景

- ユーザの意図や嗜好に合った推薦を行う技術が盛んに研究されている
- 広く利用されている推薦手法の1つに、**協調フィルタリング法**がある
  - ユーザベース方式
  - アイテムベース方式
  - バイクラスタリング法



## 問題点

- アイテムの種類は非常に多い  
→すべてのアイテムに対して嗜好が類似しているユーザが常に存在するとは考えがたい
- 一部のアイテムに対する評価が類似しているなら嗜好が類似していると仮定

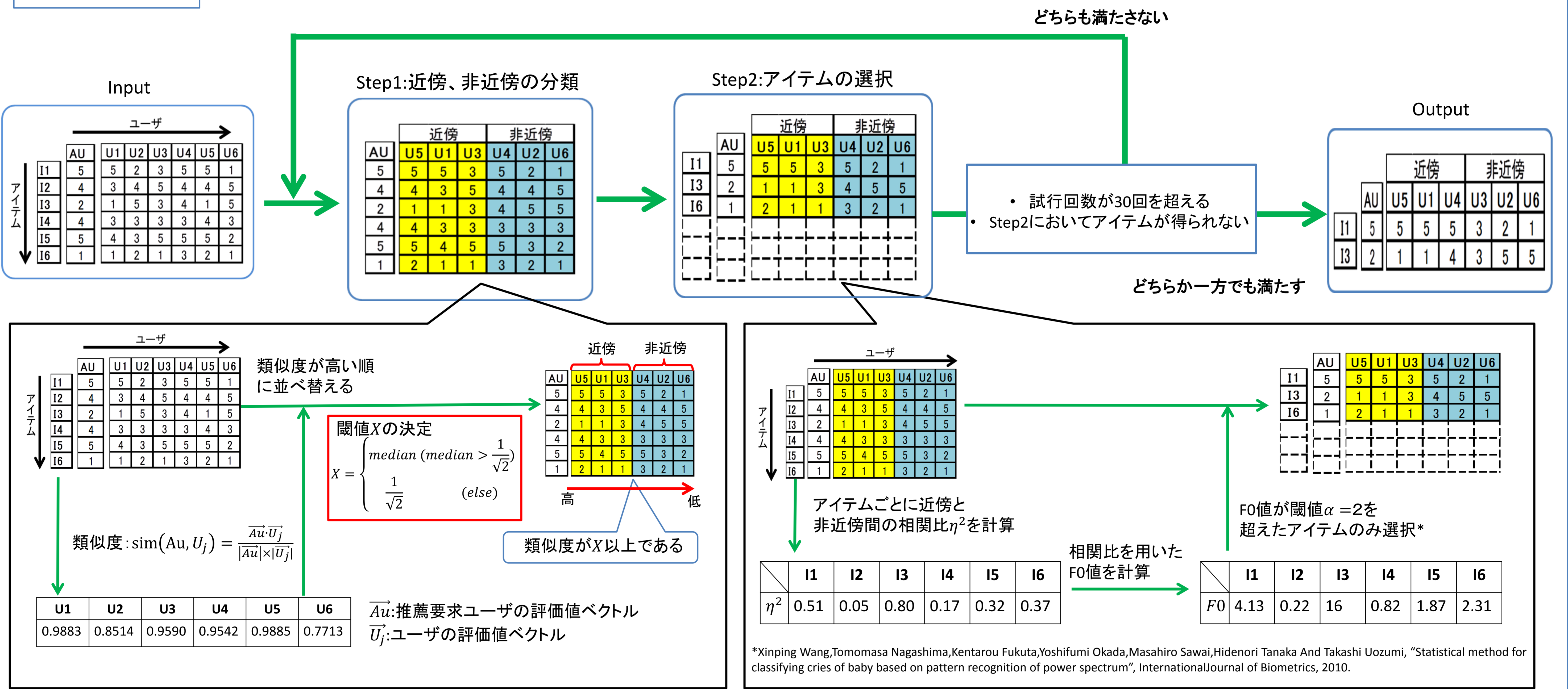


## 目的

ユーザ間で評価に差のあるアイテムのみを用いて、高精度な近傍形成を行う手法の開発

本研究では、アイテムの更新と近傍ユーザの選択を繰り返し行い、ユーザごとに近傍ユーザと非近傍ユーザを識別するアイテム集合を算出する方法を提案

## 提案手法



## 実験方法

- 毎回の更新で得られるアイテム群と近傍ユーザ群を使用し、推薦精度を算出
- 推薦精度として適合率、再現率、F1値を用いる
- 適合率 (Precision) =  $\frac{|A \cap B|}{|B|}$ , 再現率 (Recall) =  $\frac{|A \cap B|}{|A|}$
- $$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$
- A: ユーザが高評価したアイテム集合
- B: 推薦結果のうちユーザが評価したアイテム集合
- データセットとして MovieLens\* 及び Jester\* を用いる
- 推薦精度算出のため、ユーザごとに初期アイテムを設定
- 推薦スコアが上位のアイテムを推薦
- 交差検定法 (3-fold cross-validation) を使用

### データセット

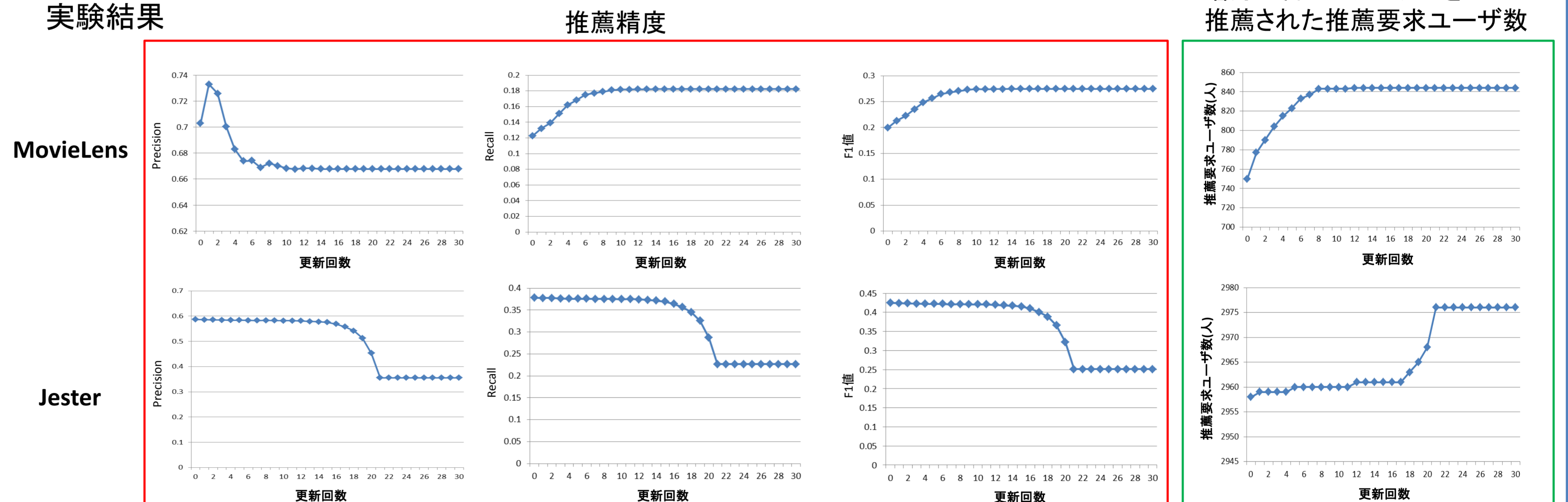
|        | MovieLens         | Jester            |
|--------|-------------------|-------------------|
| 対象データ  | 映画                | ジョーク              |
| ユーザ数   | 900               | 3000              |
| アイテム数  | 1682              | 100               |
| 最低評価件数 | 10                | 100               |
| 総評価数   | $1.0 \times 10^5$ | $3.0 \times 10^5$ |
| 評価値    | 1~5               | -10.0~10.0        |

### 実験パラメータ

|         | MovieLens | Jester |
|---------|-----------|--------|
| 初期アイテム数 | 15        | 20     |
| 推薦アイテム数 | 100       | 20     |

## 結果・考察

### 実験結果



### MovieLens

- 適合率: 初回の更新で3%程度上昇し、13回目の更新までで8%程度低下、その後収束した
- 再現率: 13回目の更新までで6%程度上昇した
- F1値: 再現率と同様の推移
- 嗜好に合ったアイテムを推薦されたユーザ数は、11回目の更新まで単調に増加し、その後収束した

### 考察

#### MovieLens

- F1値の増加 → 総合的な精度の上昇
- より多くのユーザに嗜好に合ったアイテムを推薦 → アイテムの更新が有効に働いている

### Jester

- 21回目の更新までで、推薦精度の値が減少し、その後収束した
- 嗜好に合ったアイテムを推薦されたユーザ数は17回目までは緩やかに増加し、21回目まで急激に増加した後、収束した

#### Jester

- 推薦精度の低下 → 選択されたアイテムとは異なるジャンルのアイテムに対して推薦できていない
- 嗜好に合ったアイテムを推薦できたユーザ数は18人程度増加 → 全体の人数から見るとあまり有効に働いていない

## まとめ

- 近傍と非近傍を適切に識別するアイテム群を算出し、近傍ユーザを自動的に選択する手法を提案
- アイテムの更新回数と推薦精度の関係を調査した

- MovieLensでは総合的な精度は上昇し、より多くのユーザに嗜好に合ったアイテムを推薦できた
- Jesterでは精度が大きく低下し、嗜好に合ったアイテムを推薦されたユーザ数も大きく増加していない
- 終了条件の見直し