

ネットワークモデルに基づいたエージェントシミュレーションによる情報推薦システム評価方法の提案

岡本幸樹[†] 佐賀亮介[†] 松本一教[‡]

神奈川工科大学情報学部情報工学科[†]

神奈川工科大学大学院工学研究科情報工学専攻[‡]

1. はじめに

近年、インターネット上を流れる情報は膨大なものになり、ユーザが自身の嗜好に合ったコンテンツを見つけるのは困難になりつつある。そこで、ユーザの嗜好に合ったコンテンツを探し出し提供する推薦システムが注目を集めている。様々な推薦システムが開発されているが、その多くは推薦システムの検証に人による実験を用いている。実験には再現性などの点で多くの問題がある。本研究では、エージェントシミュレーションによる推薦システムの評価方法を提案する。

2. 推薦システム

2.1 概要

推薦システムとは「ユーザの嗜好に合ったコンテンツ（以下アイテム）を探し出し提供するシステム」である。現在主に利用されている推薦方式は協調フィルタリングである。これはユーザがアイテムに付けた評価値を元に推薦する方式である。協調フィルタリングの代表的なアルゴリズムである GroupLens の評価予測値¹⁾を以下に示す。

$$p_i(b_k) = \bar{r}_i + \frac{\sum_{a_0 \in A_1} s(a_i, a_0)(r_0(b_k) - \bar{r}_0)}{\sum_{a_0 \in A_1} |s(a_i, a_0)|}$$

ここで p は予測評価値、 b はアイテム、 r は評価値、 a はユーザ、 s は類似度である。

2.2 既存研究と問題点

推薦システム開発の多くは、システムの検証に実験を用いているが、実験には以下の3つの問題点がある。

第一に環境構築の難しさである。推薦システムが利用される環境ではユーザ数は数百万、アイテム数は一千万以上になり、これだけの環境

を構築するのは容易ではない。

第二に環境を任意に設定できない点である。例えば竹川ら²⁾はユーザ数 20、アイテム数 300 で実験を行っているが、ユーザの嗜好やアイテムの種類が一様に分布しているとは限らない。

第三に再現性が低い点である。一度実験をしてしまうとユーザは知らなかったアイテムを知ってしまうので、全く同じ環境での実験は難しい。

上記の問題点を解決するために、本研究では推薦システムを検証するためのシミュレータを作成する。

3. ネットワークモデルに基づいたエージェントシミュレータ

3.1 概要・構成

本章では、提案手法を用いたシミュレータについて述べる。このシミュレータの目的は「情報推薦による情報共有をシミュレーションする」である。ユーザ、アイテム、推薦システムからなるシミュレーション環境を内部に持ち、その環境内で情報推薦をシミュレーションした後、推薦システムの評価指標を計算して出力する(図1)。

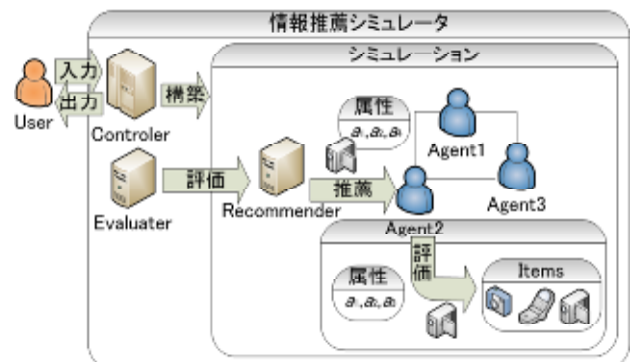


図1 シミュレータの構成

3.2 ユーザのモデル化

シミュレータでは、推薦システムのユーザのモデル化にエージェントを利用する。エージェントとは、自ら考え行動するオブジェクトを指

Proposal of evaluation method for recommender system by using agent-based simulator based on network model.

[†]Kouki Okamoto, [†]Ryosuke Saga[†], [‡]Kazuki Matsumoto

[†]Kanagawa Institute of Technology, Faculty of Information, Department of Information and Computer Sciences

[‡]Kanagawa Institute of Technology, Graduate School of Engineering, Course of Information and Computer Sciences

す。本研究のエージェントは、推薦されたアイテムを評価する機能を持つ。また、ユーザの嗜好を表す属性、アイテムの評価値、既知情報、飽き情報を持つ。

アイテムの評価モデルとしてエージェントとアイテムに属性を持たせ、お互いの属性が一致するほど、評価値が高くなるようなモデルを構築する。ここでエージェントはすべての属性に対して嗜好度という実数値を持っており、嗜好度が閾値を超えたときのみ、その属性を持つと定義する。アイテムの評価値の計算式を以下に示す。

$$S = \frac{\sum_{u_i} \frac{u_i}{\max(u_j)}}{|U \cap I|}$$

ここでSはアイテムの評価値、Uはエージェントが持つ属性、Iはアイテムが持つ属性、uは嗜好度である。

協調フィルタリングにおいて注目されている属性として「飽き」がある。シミュレータでは、エージェントに飽き指数というパラメータを持たせ、その値が閾値よりも小さいときに、飽きしていると定義する。飽き指数の計算式を以下に示す。

$$t = \sum_{i \in I} \frac{n_i R_i}{\text{rank}_i}$$

ここで、tは飽き指数、Iは推薦されたアイテム集合、rankはランキング、Rは評価値、nは既知パラメータである。

3.3 ネットワークモデル

協調フィルタリングに必要なパラメータのひとつが「ユーザ間の嗜好の類似度」である。これをモデル化するために、本研究ではネットワークモデルを利用した。各ユーザを頂点、ユーザ間類似度を辺としてモデル化し、ランダム・グラフ、スモールワールド、スケールフリーの三種類のネットワークを実装した。

3.4 推薦システムの評価

推薦システムの評価には、予測の正確さの指標と、それ以外の指標がある³⁾。本研究では、正確さの指標に MAE、recall、precision を利用し、それ以外の指標に novelty と discovery を利用する。

4. 検証実験

4.1 目的・環境・プロセス

提案手法の有用性を示すために評価実験を行う。GroupLens の協調フィルタリングを用い、「精度が向上すると飽きを発生させる」という特徴が結果に表れた場合、有用性があると判断

する。エージェント数を 100、ターン数を 10、飽き指数を 0.3、1 ターン毎の各エージェントへの推薦アイテム数を 20 とした。

4.2 実験結果及び考察

各ターンの飽きているエージェントの割合と precision と discovery をグラフに表したものを図2に示す。シミュレーションの後半で、精度の向上に伴って飽きているエージェントが増加しているのがわかる。このとき GroupLens の特徴が表れたので、提案手法の有用性が示された。

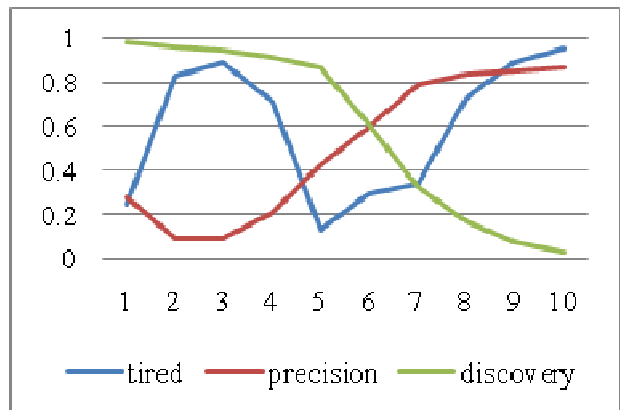


図2 実験結果

5. おわりに

本研究では、エージェントシミュレーションを利用した推薦システムの評価方法を提案した。今後の課題としては、シミュレーション結果の可視化が挙げられる。結果の可視化機能をシミュレータに組み込むことにより、各ターンのネットワーク構造や、エージェントの飽き状態等も可視化することができる。結果の可視化は入力の手軽さとともに GUI で実装したいと考えている。

[参考文献]

- 1) P.Resnick, N.Iacovou, M.Suchak, P.Bergstrom, and J.Riedl : GroupLens: Open Architecture for Collaborative Filtering of Netnews, In Conference on Computer Supported Cooperative Work, pp.175-186 (1994)
- 2) 竹川和毅, 土方嘉徳, 西田正吾 : 内容に基づく音楽データの探索・推薦システム, 人工知能学会論文誌, Vol.23, No5, pp.330-343 (2008)
- 3) 土方嘉徳 : 嗜好抽出と情報推薦技術, 情報処理学会誌, Vol.48, No.9, pp.957-965 (2007)