

閲覧サイトの登録を支援するRSSリーダー

滝沢 敏裕

山名 健悟

湯浅 将英

大山 実

Toshihiro Takizawa † Kengo Yamana † Masahide Yuasa ‡ Minoru Ohyama ‡

1. はじめに

近年 RSS が注目されており、Blog の多くは RSS を配信している。またニュースサイトでも RSS による更新情報の配信を始めている。RSS の利用には RSS リーダが必要である。しかし、多くの RSS リーダでは閲覧サイトが配信している RSS を探し出すという手間がかかる。そこで我々は利用情報を共有するリーダーを作った[1]。これはランキング表示や検索、推薦機能によって RSS の登録を支援するものである。本稿ではさらに推薦機能を発展させ、カテゴリ定義を用いた推薦方法について述べる。

2. 従来のシステムについて

本研究で開発した RSS リーダはサーバサイド型である。本システムの特徴はシステムを利用しているユーザーの RSS 利用情報を蓄積し、他のユーザーに RSS の情報をランキング、検索、推薦の機能で提供することにある。これらの情報によってボタンひとつで RSS を登録することができる。

システムを利用しているユーザーが登録している RSS を集計することでランキングを作成する。これにより、人気のある RSS やユーザーが興味のある RSS を簡単に発見することができる。

また、ユーザーが登録している RSS の記事はシステムによって収集され、検索ができる。これによりユーザーはキーワードをもとに特定の情報に関する記事を配信する RSS を見つけることができる。

さらに、ユーザーごとの RSS 登録状況を用いて協調フィルタにより、ユーザーが好むであろう RSS を推薦する。

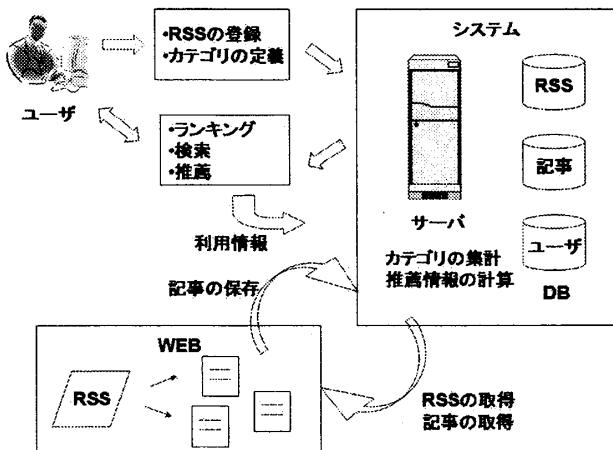


図1. システム全体の流れ

しかし、この方法ではユーザーの嗜好に合った推薦が十分ではなかった。また登録数が増えると推薦の精度が落ちてしまう。そこで、今回のシステムではカテゴリ定義を利用した推薦手法を提案する。

3. カテゴリ定義を用いた推薦

RSS にはブックマークと違い RSS 自体や記事に subject としてカテゴリが用意されているものがある。そこでこのようなカテゴリを用いて推薦の精度を上げることを考える。

3.1 カテゴリ定義

カテゴリを用いた推薦の際には二通りのカテゴリ定義を用いる。

- (1) ユーザーによるカテゴリ定義
- (2) 多数決によるカテゴリ定義

(1)はユーザーがカスタマイズページを通して RSS に自分なりのカテゴリを定義することである。図2のようにカテゴリの分類をする際、あらかじめシステムに用意されたカテゴリの中から、登録している RSS がどのカテゴリに分類されるか判断し設定する。当てはまるカテゴリがなければ、ユーザーは新たにカテゴリを追加して RSS のカテゴリ定義ができる。

(2)は RSS に対してユーザーが定義したカテゴリの集計をとり、最も多く設定されているカテゴリをその RSS のカテゴリとすることである。

これら二つのカテゴリ定義を、協調フィルタリングを用いた推薦の結果と組み合わせることで推薦精度の向上を図った。

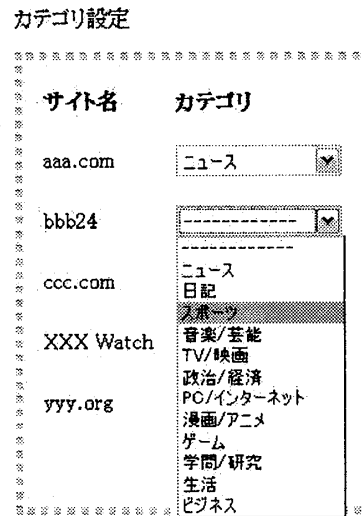


図2. カテゴリの定義のカスタマイズページ

† 東京電機大学 情報環境学研究科 情報環境工学専攻

‡ 東京電機大学 情報環境学部 情報環境工学科

ユーザ  $\alpha$

RSS	aaa.com	bbb.net	ccc
-----	---------	---------	-----

ユーザ  $\beta$

RSS	aaa.com	bbb.net	ccc	xxx	yyy	zzz
-----	---------	---------	-----	-----	-----	-----

図3. ユーザ  $\alpha$ ,  $\beta$  の RSS 登録状況

ユーザ  $\alpha$

RSS	aaa.com	bbb.net	ccc
多数決	スポーツ	スポーツ	芸能
ユーザ	スポーツ	スポーツ	芸能

ユーザ  $\beta$

RSS	aaa.com	bbb.net	ccc	xxx	yyy	zzz
多数決	スポーツ	スポーツ	芸能	芸能	日記	スポーツ
ユーザ	スポーツ	スポーツ	音楽	芸能	日記	なし

図4. それぞれのカテゴリ定義

### 3.2 推薦手法

図3に従来までの協調フィルタを用いた例を示す。図ではユーザ  $\alpha$  と  $\beta$  の RSS (aaa.com~zzz) の登録状況が示されている。登録状況の類似度計算によりユーザ  $\alpha$  と  $\beta$  が近いとされたとき、ユーザ  $\alpha$  には RSS xxx, yyy, zzz が推薦される。

今回のシステムでは従来の類似度計算に加えて新たにカテゴリによるフィルタ処理をする。フィルタを用いて考えられる推薦手法は以下である。

1. 従来の手法
2. 推薦対象  $\alpha$  : ユーザ 推薦者  $\beta$  : 多数決
3. 推薦対象  $\alpha$  : ユーザ 推薦者  $\beta$  : ユーザ
4. 推薦対象  $\alpha$  : 多数決 推薦者  $\beta$  : 多数決
5. 推薦対象  $\alpha$  : 多数決 推薦者  $\beta$  : ユーザ  
(ユーザ, 多数決はそれぞれカテゴリ定義の方法)

たとえば、図4のような状況で推薦手法4を用いた場合では、まず推薦対象であるユーザ  $\alpha$  が登録している RSS の多数決によるカテゴリ定義を調べる。ユーザ  $\alpha$  の登録している RSS のカテゴリはスポーツ、芸能である。次に推薦者  $\beta$  から推薦される RSS における多数決によるカテゴリ定義は、それぞれ (xxx, 芸能) (yyy, 日記) (zzz, スポーツ) である。ここでユーザ  $\alpha$  が登録している RSS カテゴリと比較し、xxx と zzz を推薦する。ユーザ  $\alpha$  のもっていないカテゴリである yyy の日記は推薦されない。

表1. 手法毎の採用率

推薦手法	採用率
推薦1	60.0%
推薦2	72.2%
推薦3	72.2%
推薦4	75.0%
推薦5	75.0%

表2. ユーザ毎の採用率

ユーザ	採用率
A	93.7%
B	93.7%
C	100%
D	50.0%
E	100%
F	46.6%
G	0.0%
H	78.5%

### 3.3 実験

5種類の推薦手法を比較するため実験をした。理系学生8人のユーザ (A~H) に自分の嗜好にあった RSS を最低4つ~5つ登録してもらい、登録した RSS についてカテゴリの設定をしてもらった。それぞれの推薦結果を見て推薦された RSS が自分の嗜好に合っており、登録するか否かを答えてもらった。表1, 2はそれぞれ推薦の採用率 (登録の可・否) である。

### 3.4 考察

推薦手法1ではカテゴリによるフィルタ処理をしていないため、ユーザが興味のないカテゴリが含まれていても推薦している。そのため他の推薦よりも若干低く60%の推薦採用率となった。

推薦手法2, 3は72.2%と同じ値になっている。さらにユーザ毎の詳細を調べると、推薦手法2と3では同じ結果になる傾向があった。推薦手法3ではあらかじめ用意されたカテゴリよりも細かいカテゴリをユーザが追加することでユーザ間の共通点を見つけ出すことを期待したが、今回の実験では人数が少なかったため変化が現れなかった。

ユーザごとの推薦採用率ではユーザ G には嗜好にあった推薦ができなかった。ユーザ G は RSS リーダ利用者の中で最も RSS の登録数が多かったユーザであった。ユーザ G は推薦された RSS のカテゴリ自体に興味があるが、既に同じカテゴリに属する別の RSS を登録しているため、推薦された RSS は必要ないと判断した。

ユーザ D, F の推薦採用率が低い理由として、今回の実験をしたユーザとは極めて異なる嗜好を持っているため、採用率が低くなってしまった。これはユーザ数を増やすことで改善できると思われる。

### 3.5 新たな推薦手法の提案

今回の実験から登録状況に対して協調フィルタリングを用いた類似度の計算をすると、ユーザ G のように RSS の登録数が多いユーザでは推薦の精度が落ちてしまう。そこで、ユーザが登録している RSS のカテゴリ毎に類似度の計算をすることで、カテゴリ毎に最も似ているユーザの RSS を推薦することも考えられる。これにより似ているユーザを一人見つけ出す現在の推薦よりも精度の高い推薦が可能であると考えられる。

### 4. まとめ

ユーザの利用情報を用いて閲覧サイトの RSS 登録を支援する RSS リーダを開発した。ユーザによるカテゴリ定義と多数決によるカテゴリ定義を用いた推薦手法を提案し、実装及び実験を行なった。単純な協調フィルタリングによる推薦よりも良い推薦精度となった。

今後は、ユーザ数を増やし推薦結果の変化について調べることで、新しい推薦方法について実装と実験をする予定である。

### 参考文献

[1] 滝沢, 山名, 湯浅, 大山: “複数ユーザの利用情報を用いた RSS リーダ”, 情処学, 第67回全国大会, 3, pp.197-198(2005)