

# SNS の情報を利用した RSS フィルタシステムの設計と実装

券田 孝晴<sup>†</sup>西山 裕之<sup>†</sup><sup>†</sup> 東京理科大学大学院 理工学研究科

## 1 はじめに

RDF Site Summary (以下 RSS) とは、ニュースサイトやブログなど、Web サイトの見出しや内容などの情報を XML 形式でまとめたものである。RSSリーダーという専用のソフトを使うことで、登録した RSS のニュース記事、ブログ記事の更新情報を容易に得られる。しかし最近では、大量に更新する Web サイトの RSS の存在や、RSS 自動生成システムのあるブログの流行などによる RSS の数自体の増加によって、ユーザの取得情報量が膨大になり、ユーザが興味を持ち得る情報を選別することが困難になっている。

一方、SocialNetworkingService(以下 SNS) という、人と人とのつながりを形成、サポート、推進することを目的しているサービスが注目されている。SNS のような社会的ネットワーク上での行動データとそのユーザの特徴の間に相関関係があるという結果を出した研究 [3] がある。この結果から、ユーザの興味がある情報は、ユーザと近い特徴を持つ友人も興味を持ち得る情報であると考えられる。当然、逆も然りである。

本研究ではユーザの効率的な情報収集活動を支援することを目的とした、SNS の情報を利用した RSS フィルタシステムの実装を行う。ユーザの特徴を示しているユーザの情報 (閲覧履歴、評価、RSS データ、ユーザキーワード) だけでなく、近い友人の情報も利用し、ユーザが優先的に興味を持ち得る記事を選択できるようなフィルタを、情報 (協調) フィルタリング [1] の手法を適用することで構築する。本システムを利用しフィルタ後の記事を読むだけで効率的な情報収集が可能となり、既存の RSS の問題点を解消する。なお、本システムは、Web 上で動作し、ユーザーに関する情報収集と情報フィルタリングから構成され、同時に、そのフィルタされた記事情報をユーザーに提示するインターフェースを備えている。本システムの概要を図 1 に示し、第 2,3,4 章で本システムの設計について述べる。

## 2 RSS 表示インターフェース

ユーザに対しての記事情報の提示や、ユーザの様々な情報を取得するためのインターフェースについて述べる。

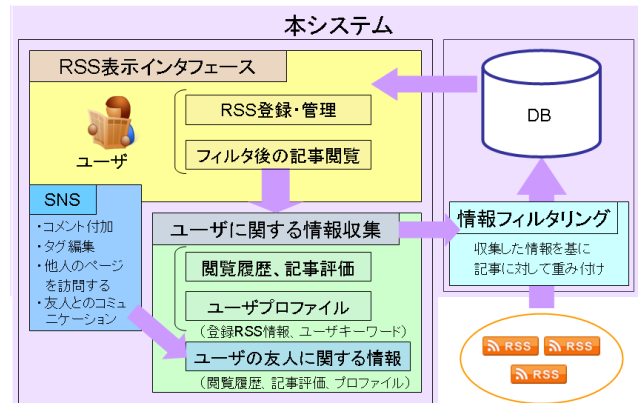


図 1: システム概要

### RSS 管理・記事閲覧

RSS 管理では RSS をフォルダごとに整理するように管理が可能であり、フォルダごとのフィルタリングも可能である。記事閲覧ではフィルタ後の記事をランキング表示形式や記事表示形式 (図 2 右に記事表示ページを示す) でユーザに提示し、ユーザの情報収集を支援する。

### SNS

ソーシャルネットワーク構築のために、ユーザ情報管理機能としてプロフィール機能からユーザの詳細な情報を記録する。そして、コミュニケーション機能として、記事へのコメントやメッセージ送信が可能である。また記事にタグを自由に付けられ (元々フィードに記されているカテゴリとは別としている)、そのタグから記事をソートする機能がある。



図 2: 左: ユーザー TOP, 右: 記事表示ページ

The RSS filter system which used information of SNS

Takaharu Kenda<sup>†</sup>, Hiroyuki Nishiyama<sup>†</sup><sup>†</sup> Graduate School of Science and Technology Tokyo University of Science

### 3 ユーザに関する情報収集

個々のユーザに対して効率的な情報フィルタリングを行うために、各ユーザの情報として閲覧・評価情報とユーザプロフィール (RSS データ, ユーザキーワード) 情報について述べる。

#### ユーザの情報

ユーザ情報は、ユーザ自身が入力した RSS データと RSS 自体の評価値 (1 から 10 の整数値), ユーザキーワードとその評価値, 記事を読んだときの評価, そして, SNS 内での行動データ, デモグラフィックデータからなる。ユーザキーワードは、記事に評価を付けた時の評価値 (1 から 5 の整数値) によって、記事のタイトル及び概要に形態素解析を行い、キーワードを TF-IDF 法に従い重みをつけてユーザプロフィールに追記される。

#### 友人の情報

SNS 内の友人の中でもよりユーザと近い友人の情報のほうが有効であると考えられる。そこで本研究では、友人間の類似度を Breese ら [2] の協調フィルタリング手法を応用して測り、フィルタに活用する。ユーザ  $a$  が持つある一つの RSS 内で、ユーザ  $a$  が評価した記事  $j \in l_a$  ( $l_a$  はユーザ  $a$  が評価した記事集合) の評価を  $v_{aj}$  と表すと、ユーザ  $a, i$  が共通して持つある一つの RSS での、ユーザ  $a, i$  間の類似度  $w_{ai}$  は

$$w_{ai} = \frac{\sum_{j \in l_i} (v_{aj} \times v_{ij})}{\sqrt{\sum_{k \in l_a} v_{ak}} \sqrt{\sum_{k \in l_i} v_{ik}}} \quad (1)$$

となる。さらに、ユーザ  $a$  が持つ RSS 集合を  $r_a$ , ユーザ  $a$  が持つある一つの RSS  $n \in r_a$  の評価値を  $h_{an}$  と表すと、ユーザ  $a$  のユーザ  $i$  に対する類似度  $S_{ai}$  は次式で定義される。

$$S_{ai} = \frac{\sum_{n \in r_a \cap r_i} (w_{ain} \times h_{an})}{\max_{n \in r_a} h_{an} \times |r_a|} \quad (2)$$

(2) 式は、ユーザ  $a, i$  が共通して持つある一つの RSS での、ユーザ  $a, i$  間の類似度の重みを、ユーザ  $a$  の RSS 評価値で付け、それがユーザ  $a$  が持っている全ての RSS の中でどのくらい重要なのかを計算した式である。

### 4 情報フィルタリング

第 3 章で集めた、ユーザの情報を基に、RSS からの最新記事に情報フィルタリングを行う際の手法について述べる。ユーザ  $i$  の記事評価の平均値  $\bar{v}_i$  は

$$\bar{v}_i = \frac{1}{|L_i|} \sum_{n \in r_i} \sum_{j \in l_i} v_{ij} \quad (3)$$

となる。(3) 式の  $L_i$  は、ユーザ  $i$  が評価した全ての記事集合である。(3) 式を用いて、ユーザ  $a$  に対する記事  $j$  の有用値  $P_{aj}$  は次式で表すことができる。

$$P_{aj} = \bar{v}_a + \frac{\sum_{i \in U} (S_{ai} \times (v_{ij} - \bar{v}_i))}{\sum_{i \in U} S_{ai}} \quad (4)$$

(3)(4) 式で用いる記事  $j$  は、ある一つの RSS  $n$  内の記事の一つとしている。また、(4) 式の第 2 項は、ユーザ  $i$  が記事  $j$  に評価した値から、ユーザ  $i$  の記事評価の平均値を減算したものを、類似度  $S_{ai}$  で重みを付け、ユーザ  $a$  以外の他のユーザ全員の対して総和をとり正規化したものである。さらに、 $P_{aj}$  はユーザプロフィール (RSS データ, ユーザキーワード) 情報の評価値を加算して再計算される。本研究では、この値に基づく記事の順位付けを行う。値は 1 日 1 度の自動更新で計算される。

### 5 考察

本システムと従来の Web 型 RSS リーダーとの比較を行い、本システムの優位性を考察する。ここでは、代表的な RSS リーダーとして、様々な機能を持つ goo RSS リーダーと Google Reader との比較を行う。比較結果を表 1 に示す。

表 1: 従来の Web 型 RSS リーダーとの比較

	goo リーダー	GoogleReader	本システム
RSS リーダー機能			
キーワード フィルタ機能			
記事評価機能			
SNS 的機能	×		
RSS 自体の評価機能	×	×	
ユーザに特化した 記事のランキング表示	×	×	

( : 完全に対応している, : 一部対応している, ×: 対応していない)

表 1 より、本システムでは、従来の Web 型 RSS リーダーにはない機能により、現状の RSS の問題点を解決していることがわかる。さらに、本システムは、ユーザが記事に付加情報 (タグ, コメント) を自由につけることができるので、ユーザにランキング以外の情報選別フィルタを提供することが可能である。

### 6 おわりに

本研究では、ユーザが優先的に興味を持ち得る記事を選択できる SNS の情報を利用した RSS フィルタシステムの設計と実装を行った。また、従来の Web 型 RSS リーダーと比較して本システムの優位性を示した。本システムを用いることで、既存の RSS の問題点を解消する、ユーザの効率的な情報収集支援を実現した。

#### 参考文献

- [1] 森田昌宏, 速水治夫: 情報フィルタリングシステム, 情報処理, Vol.37, No.8, pp.751-758(1996).
- [2] Breese, J.S, Heckerman, D. and Kadie, C.: Empirical Analysis of Predictive Algorithms for Collaborative Filtering, Proc. 14th Conference on Uncertainty in Artificial Intelligence, pp.44-52(1998).
- [3] Singla, P., Richardson, M.: Yes, There is a Correlation: From Social Networks to Personal Behavior on the Web. In: Proceeding of the 17th International Conference on World Wide Web, Apr. 21-25(2008).