

書籍情報から得た個人情報の ソーシャルネットワークへの活用

塚本 哲也[†]西山 裕之[†]大和田 勇人[†][†]東京理科大学大学院 理工学研究科

1 はじめに

Social Networking Service(以下 SNS) は人と人とのつながりを形成, サポートすることを推進しているサービスである. SNS は世界中で広がりつつあり, 日本国内最大級の SNS 「mixi」 では 600 万人を超えるユーザーが参加しているという調査結果 [1] がある. 友人や知人同士のコミュニケーションの場や趣味趣向の合う他人とのやり取りの場を提供しているが, SNS には趣向が合う他人を探す手段が乏しい. ユーザーの趣向を表現する情報が容易には手に入らず, 個人の行動データを利用して他のユーザーとマッチングするような機能が少ないことが問題である. また, 社会的ネットワーク上での行動データとそのユーザーの特徴の間に相関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ーの特徴の指標に蔵書情報を利用した場合に対しても, ソーシャルネットワークとユーザーとの間に相関関係があるのかをデータマイニングにより明らかにすることは有用であると考えられる.

本論文では SNS と蔵書管理システムの特長を融合したシステム「ItemSpider[2]」を用いて, ユーザーの個人情報に蔵書の構成を属性として取り入れ, ユーザーの類似性を計測することが目的である. ユーザーの蔵書情報というパーソナルな情報を加えることで SNS 内のユーザーの特徴を詳細に表現することができるようになる. それを元に他ユーザーとの比較や, データマイニング手法を適用して新たなコミュニティの発見を行うことで SNS の問題点を解消する. その結果, より趣味が合うユーザーを発見することを目指す. 本システムの機能はユーザーに関する情報収集とユーザーの特徴付加の 2 部で構成される. 本システムの概要を図 1 に示す.

2 ユーザに関する情報収集

情報収集部ではユーザーの特徴を計算するために必要なユーザーの振る舞いを SNS, ユーザーの蔵書から集計すると同時に, その情報をユーザーに公開するインターフェース (図 2 左に ItemSpider のユーザー TOP, 図 2 右に書籍個別ページを示す) を備える. ItemSpider の詳細な実装項目については先行研究 [2] を参照されたい. 以降, ItemSpider を通じて取得するデータと閲覧インターフェースについて述べる.

Social Network Service that applies personal character extracted from book collection.

Tetsuya Tsukamoto[†], Hiroyuki Nishiyama[†], Hayato Ohwada[†]

[†]Graduate School of Science and Technology Tokyo University of Science

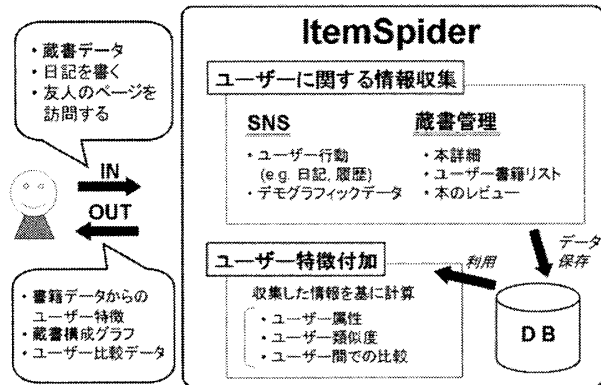


図 1: システム概要

2.1 SNS

ソーシャルネットワーク構築のためのユーザー情報管理機能と, コミュニケーション機能を実装し, ユーザーの特徴と行動を取得するため利用履歴を取る. ユーザー情報管理機能では, ログイン処理モジュールと行動ログモジュールからユーザー行動ログ, プロフィール機能からユーザー情報を記録する. コミュニケーション機能では, 日記やコミュニティへのコメントやメッセージ履歴からユーザーの行動ログを取得する.

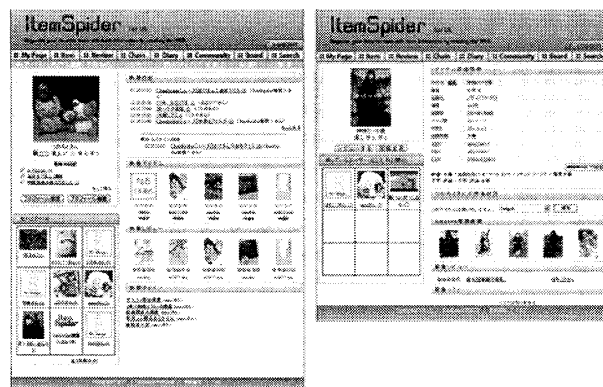


図 2: 左: ユーザー TOP, 右: 書籍ページ

2.2 書籍情報

ユーザーは ISBN またはキーワードを用いた検索から詳細な書籍データをシステムに追加できる. 書籍の

情報は Web API (Amazon Web Service) を通じて自動取得する情報 (表 1) と、ユーザーが独自に書籍へ付加する情報がある。図 3 にユーザーが登録した書籍情報を閲覧できる蔵書ビューを示す。

表 1: 書籍データ詳細 (※が付いたものは複数項目)

ISBN-10	ISBN-13	ASIN	書籍タイトル
※著者	※著者分類	出版社	※推薦書籍
出版日	出版形態	価格	サイズ
書籍表紙イメージ		Amazon 個別 URL	
Amazon 商品ジャンル		※ジャンル情報	

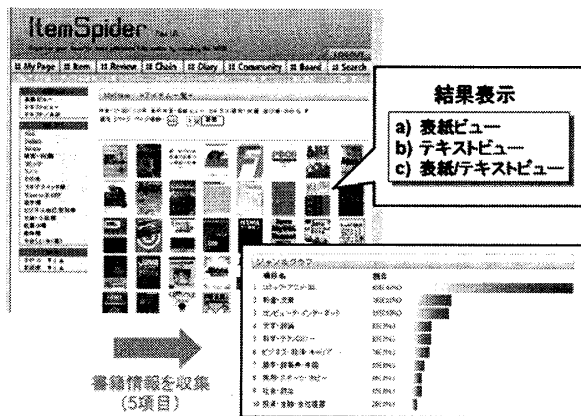


図 3: 蔵書ビューと蔵書構成グラフ

3 ユーザー特徴付加

ユーザーに対して特徴を付けるために、本研究では ItemSpider を通じて SNS 上での振る舞いと各ユーザーの書籍情報を取得している。まず、ItemSpider が収集した情報を利用したユーザー属性と、ユーザー間類似度を計測する手法の定義を行う。

ユーザー属性

ユーザー属性値は、5 項目 (書籍ジャンル、著者、出版社、出版日、価格帯) の要素とその出現数からなるベクトル 5 種、SNS 内での行動データ、ユーザー自身が入力したデモグラフィックデータからなる。本論文では蔵書データから得た 5 項目をユーザー属性として扱う。詳細な内容は表 2 にまとめた。

表 2: ユーザー属性ベクトル詳細

ジャンル	27 種の Amazon ジャンルコード
著者	全ての著者名
出版社	全ての出版社名
出版日	月ごとに再分類した出版日 (yyyy-mm)
価格帯	本の価格を 6 分割し再分類した価格

ユーザー間類似度

ユーザー間の類似度はベクトル空間法を用いて求める。ユーザー属性は 5 つのベクトルになっており、各ベクトルごとにベクトル空間法を適用する。

3.1 ユーザー間の特徴比較

ユーザー間の特徴を比較する手順に沿って説明する。まずシステムは、全ユーザーのユーザー属性値を取得し、全てのユーザーの組み合わせについてユーザー間類似度を計算する。そしてユーザーごとに他のユーザー全員に対して類似度の最大値、最小値をそれぞれ上限、下限に 10 等分し、類似度に応じた 0~10 のスコアをユーザーに割り振り合計を算出する。このスコア計算は 1 日 1 度自動更新される。

ユーザーは、ユーザー類似度に関するチャートグラフと詳細データ、そしてその対象ユーザーとの共通本の情報が含まれたビュー (図 4) を通じて閲覧する。これにより蔵書の特徴として類似ユーザーを一覧で検索、確認することが可能になった。

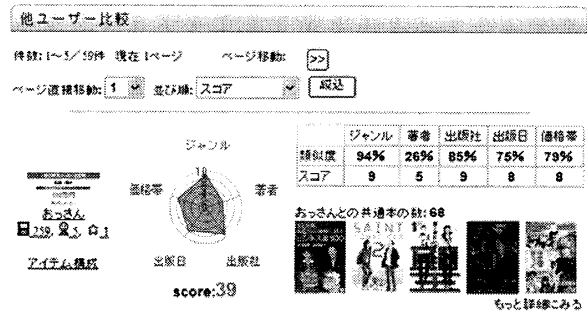


図 4: 特徴比較結果ビュー

4 おわりに

本論文では、各ユーザーに蔵書の構成をユーザーの特徴として取り入れ、ユーザー間の類似性を計測するため機能について述べた。既存の SNS と比較してユーザーの特徴、ユーザー間の類似度へアクセスが容易である。今後はユーザーの特徴をコミュニティに対して適用する。どのコミュニティがユーザーに適合しているか、コミュニティ内メンバーと類似しているかといったコミュニティ単位での特徴を明らかにする。

参考文献

- [1] 総務省情報通信白書平成 20 年度版, <http://www.johotsusintokei.soumu.go.jp/whitepaper/>
- [2] Tetsuya Tsukamoto, Hiroyuki Nishiyama and Hayato Ohwada.: ItemSpider: Social Networking Service that extract personal character from individual's book information. The Pacific Rim Knowledge Acquisition Workshop, Dec. 15-16 (2008)
- [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)