

振る舞い協調フィルタリングによるソーシャルストリームの効率的な閲覧手法

顔 洪[†] 牛尼 剛聡[‡]

[†]九州大学芸術工学府 〒815-8540 福岡市南区塩原 4-9-1

[‡]九州大学大学院芸術工学研究院 〒815-8540 福岡市南区塩原 4-9-1

E-mail: [†]3DS13013N@s.kyushu-u.ac.jp, [‡]ushiama@design.kyushu-u.ac.jp

あらまし 近年、ソーシャル・ネットワーキング・サービス(SNS)が普及し、刻一刻と新しいコンテンツが生まれている。しかし、SNS から配信される膨大なコンテンツのストリームを閲覧する場合、価値があるコンテンツを見逃す可能性があり、効果的な閲覧は容易ではない。本研究では、情報推薦の典型的な手法である協調フィルタリングを、閲覧者間の振る舞いに活用することで、効率的なソーシャルコンテンツのブラウジングを実現する手法を提案する。本手法では、対象ユーザと振る舞いが似ているユーザが、コンテンツを読んだ際の振る舞いに基づいて、対象ユーザの未読コンテンツに対する適切な提示時間を推定することにより、効率的な閲覧を実現する。

キーワード スマートフォン, SNS, ストリームコンテンツ, 協調フィルタリング, 振る舞い, 読み飛ばし

Hong YAN[†] Taketoshi USHIAMA[‡]

[†]Graduate School of Design, Kyushu University 〒815-8540 4-9-1 Shiobaru, Minamiku, Fukuoka, JAPAN

[‡]Faculty of Design, Kyushu University 〒815-8540 4-9-1 Shiobaru, Minamiku, Fukuoka, JAPAN

E-mail: [†]3DS13013N@s.kyushu-u.ac.jp, [‡]ushiama@design.kyushu-u.ac.jp

Abstract In recent years, Social Networking Services (SNSs) are growing in popularity, and generating fresh contents moment by moment. However, when huge content streams delivered from the SNS, the efficient browsing is not easy because there are serious possibility for overlook valuable contents. In this study, we propose a method to achieve efficient browsing by exploiting the behaviors of readers based on the collaborative filtering which is a typical technique of information recommendation. Based on the behaviors of the users who are similar to the target user's behavior with reading contents, our system estimates the appropriate presenting time for unread contents of the target user to realize efficient browsing.

Keyword Smartphone, SNS, stream contents, collaborative filtering, behaviors, skipping.

1. はじめに

近年、Twitter¹, Weibo²などを代表とするソーシャル・ネットワーキング・サービス(Social Networking Services)が普及している。SNS では、ユーザの経験、感想、興味等が含まれている実生活に関する様々な情報をリアルタイムに配信・記載するために利用されている。このような状況を背景に、膨大なコンテンツストリームの中から、価値があるコンテンツを効果的に発見・利用するためのストリームデータの処理が注目されている^[1]。従来、莫大なデータ空間から必要な情報を取得するために Google³ や Yahoo!⁴ 等のキーワード型の検索エンジンを用いて、ユーザにとって価値が

高いと推定される情報のみを提供してきた。しかし、Twitter や Weibo の場合は、ユーザに提供されるストリームコンテンツには、ユーザにとって価値が高いコンテンツと価値が低いコンテンツが混在しており、また、そのコンテンツの主題も様々である。SNS のユーザは、このようなストリームとして提供される大量なコンテンツをすべて閲覧するのは、ユーザに対する負担が大きい。一般に、SNS のユーザは、ストリームとして配信されるコンテンツのすべての内容を厳密に理解するわけではなく、「飛ばし読み」によって、ユーザにとって価値が高いと思われるコンテンツを見つけ、特定のコンテンツのみ内容をしっかりと確認する。しかし、「飛ばし読み」をした場合、価値があるコンテンツを見逃す可能性がある。これまでに、テキストデータを対象に、自然言語処理技術^[2]を用いて、ユーザにとって価値が高いコンテンツを推定するための研究が行われている。しかし、Twitter や Weibo のコンテンツにも、

¹ <http://www.twitter.com>

² <http://www.Weibo.com>

³ <http://www.google.com>

⁴ <http://www.yahoo.co.jp>

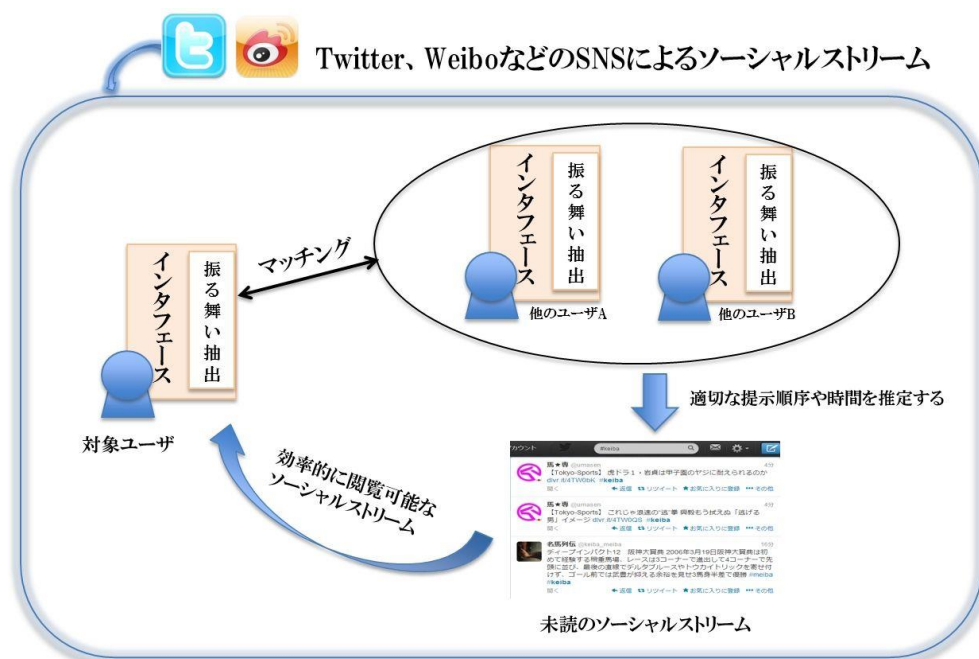


図 1. 提案手法のアプローチ

テキストデータは含まれるが、文字数が制限されているため、効果的な推定が困難であるという問題がある。

一方、スマートフォンにおけるタッチパネルを用いたインタフェースの進歩、高性能化および低価格化に伴い、スマートフォン上で日常的に SNS を利用するようになった。また、スマートフォンを利用することにより、ユーザに対して振る舞い・操作を取得できるようになった。ここで、人の振る舞いとは、スマートフォンを操作する際の指の動きあるいは手の動き等である。

本稿では、対象閲覧者に対して、振る舞いの情報によって、閲覧したコンテンツの価値があるかどうかを反映されるという仮説の下で、情報推薦の典型的な手法である協調フィルタリングの考え方に基づいて、閲覧者間の振る舞いの類似性に基づいて、ソーシャルとリームの効率的なブラウジングを実現する手法を提案する。本手法のアプローチを図 1 に示す。本手法では、対象ユーザと振る舞いが似ているユーザがコンテンツを読んだ際の振る舞いに基づいて、対象ユーザの未読コンテンツに対する適切な提示時間を推定することにより、効率な閲覧を実現する。

以下では、2 章で関連研究について述べ、3 章で本研究の位置付けおよびアプローチを紹介する。その後、4 章でユーザの振る舞いとその際に読んでいたコンテンツの価値の関係性の調査について述べる。5 章と 6 章で振る舞い協調フィルタリングに基づいたコンテンツの価値の推測手法および評価実験について説明する。最後に、7 章で本稿のまとめを行う。

2. 関連研究

2.1 情報推薦におけるフィードバックの活用

これまでに、蓄積された膨大な情報空間の中から、ユーザにとって価値があるコンテンツを発見し^[3]、推薦する研究が活発に行われてきた。それらの推薦システム^[4-7]において、最も多く利用されているのは、典型的な協調フィルタリング手法である。協調フィルタリングは、映画や音楽などのオンラインショッピングでの推薦においてよく用いられている。協調フィルタリングの基本的な考え方は、過去の行動データに基づいて、ユーザ間の類似度を介し、対象ユーザの嗜好を推測することにより、推薦するアイテムを決定する。また、いくつかの推薦システムでは、推薦した結果について、ユーザからのフィードバックを収集し、フィードバック^[8]に基づいてリアルタイムに推薦内容を最適化する。つまり、これらの手法では、ユーザのフィードバックに従って推薦結果の精度を向上させる。しかし、ユーザにフィードバックを要求することにより、ユーザの負担が増えてしまうという問題がある。

2.2 個人化ための協調フィルタリング

協調フィルタリングには様々な手法が提案されている。その中でも、モデルベースの協調フィルタリング手法とメモリベースの協調フィルタリング手法が幅広く利用されている。

● モデルベースの協調フィルタリング

モデルベースの協調フィルタリング手法は、ユ

ーザの過去の行動履歴や評価に基づいて、統計アルゴリズムを用いてユーザの行動をモデル化し、モデルに基づいて将来か価値を予測することにより、推薦対象を決定する。対象ユーザの評価を予測するためにヒューリスティックな方法の利用により、機械学習を使用しなくても予測できるモデルとなる。古典的なモデルベースの協調フィルタリング手法は、ベイジアンネットワーク (Bayesian network)^[18]、クラスタリング (Clustering)^{[19][20]}、ルールベースモデル (Rule-based model)^[25] などがある。モデルベースの協調フィルタリングアルゴリズム^[17]は、種々の典型的な統計方法を用いてノイズデータの除去を行うことによって、推薦の質を向上可能である。しかし、モデルベースの協調フィルタリングを構築するために多くの時間を要するため、リアルタイムに更新するオンライン環境の場合、速い応答には適することが困難である。

● メモリベースの協調フィルタリング

メモリベースのアプローチ^[10]は、モデルベースのアプローチよりも最初に出現した。メモリベースのアプローチは、アマゾン⁵などの商用サービスで最も広く使用されている協調フィルタリング^[21-23]の方法である。一般的に、メモリベースのアプローチは、ユーザ間の類似度に基づいて、他のユーザから評価したかつ対象ユーザが評価されていない項目の価値を予測する。最後に、トップ N 個項目を対象ユーザに推薦する。メインのユーザ類似度計算方法は、古典的なコサイン類似度 (Cosine-based Similarity^{[27][28]}) およびピアソン相関係数 (Pearson Correlation Coefficient^[29]) などがある。

対象となる個々のコンテンツやユーザの振る舞いについての詳しい分析を必要とするため、対象となるコンテンツの価値やそれらのコンテンツに対するユーザの振る舞いを推測すること^[13-16]はそれほど容易ではない。さらに、コンテンツストリームを適時に対処するため処理速度も考慮すべきである。

3. 提案手法の概要

本研究では、SNS においてユーザがスマートフォンを利用する状況を想定し、ユーザが未読のコンテンツを閲覧する際に、コンテンツストリームを適切に処理可能な閲覧手法の実現を目指す。

提案手法では、ソーシャルストリームの効率的なブラウジングを実現するために、情報推薦の典型的な手法

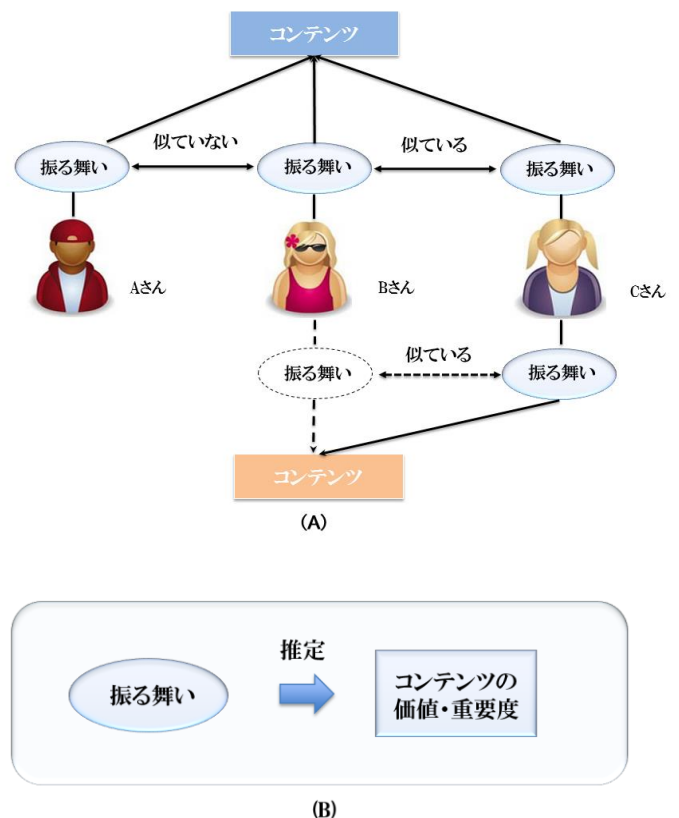


図 2.提案手法のイメージ

である協調フィルタリングを、閲覧者間の振る舞いに活用する。例えば、図 2(A)のように、対象ユーザ(ユーザ A とする)の、あるコンテンツに対する価値を推定することを考える。このとき、ユーザ A の情報に対する振る舞いを観測し、過去に、同じコンテンツに対して、ユーザ A と振る舞いの類似性が高い他のユーザ(個々のコンテンツに対する相対的な振る舞いなどがユーザ A の振る舞いと高い度合いで一致するユーザ(図 2 の場合は、ユーザ B))をコンテンツに対する価値の評価が似ていると推定する。そこで、ユーザ B が閲覧したが、ユーザ A が未読であるコンテンツの価値を推測する。この枠組みでは、各ユーザが振る舞いの情報収集することにより、質の高い推定が可能となる。

4. SNS 閲覧時のユーザの振る舞いの分析

上記のように、ユーザの振る舞いに基づいて、コンテンツの価値・重要度を推定するためには、図 2(B)に示すように、対象閲覧者に対して、閲覧したコンテンツに対する振る舞いが、そのコンテンツに対するユーザの価値判断を反映していることが重要である。

本章では、Twitter と Weibo を利用して、ユーザ振る舞いとユーザの価値判断の関連性に関する実験の結果を示す。本実験の目的は、SNS におけるコンテンツ

⁵ www.amazon.com

の価値とユーザの閲覧際の振る舞いの関係性を明らかにすることである。

コンテンツの価値とユーザの閲覧時の振る舞いの関連性を調査するため、現実の SNS のストリームデータを利用して、実験用のデータセットを作成し、データセット中のコンテンツの価値をユーザに 5 点評価してもらうことで、閲覧際の振る舞いの関係性を分析した。

4.1 実験環境

本実験では、スマートフォンを利用する。実験のために、スマートフォン上で簡潔な操作を記録するインタフェース（図 3）を開発した。このインタフェースでは、コンテンツが 1 件ずつ表示され、ユーザはフリック操作で、次のコンテンツを表示することができる。なお、使用したスマートフォンの CPU は 1.6GHz であり、メモリは 2 GB であった。なお、表示用のディスプレイサイズは、4.8 インチである。OS は Android™ 4.1.1 である。

4.2 読み速度とコンテンツの価値の関係性

コンテンツの価値と振る舞いの関係性の分析を行うため、Weibo において 2013 年 11 月 30 日に配信されたコンテンツをランダムに選択し、100 個のコンテンツをデータセットと設定した。

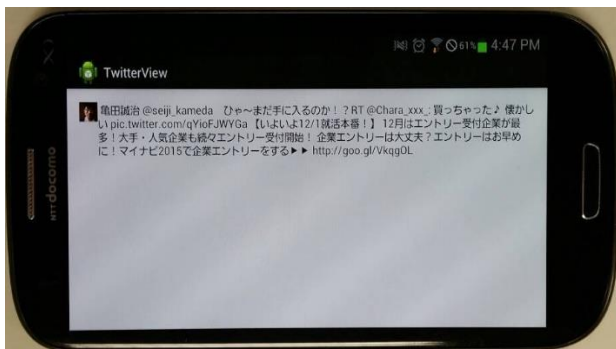


図 3. 調査実験を行うためのインタフェース

本実験では、被験者に対して、スマートフォンでデータセットのコンテンツを個々に読んでもらった。閲覧後、被験者は各コンテンツに関する価値を 1 点（最も価値が低い）から 5 点（最も価値が高い）まで評価した。一方、個々のコンテンツを読む時間を計測し、コンテンツを読むのに要する閲覧時間と価値の関係性を分析した。被験者は、日常的に Weibo を使用する大学生および大学院生の中国人 4 名である。

分析のために、コンテンツ i において、位置ミリ秒あたりに読む文字数を読み速度 rv_i として、以下の式で定義する。

$$rv_i = \frac{length_i}{time_i} \dots (1)$$

ここで、 $length_i$ はコンテンツ i の文字数であり、 $time_i$ はコンテンツ i の表示時間を表す。読み速度が大きいほど、ユーザはそのコンテンツを飛ばし読みしている可能性が高い。

各被験者によるコンテンツに対する読み速度と、そのコンテンツに対する評価を獲得し、コンテンツの価値の評価と、各被験者の読むための平均速度の関係を図 4 に示す。

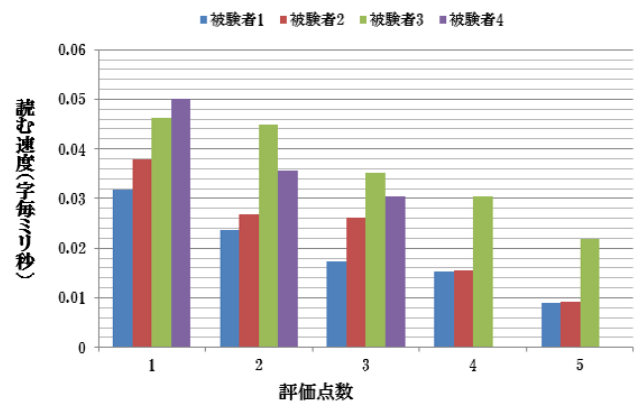


図 4. 実験による各被験者の平均読み速度

4.3 フリック速度とコンテンツの価値の関係性

ユーザが次のコンテンツを表示するためのフリックの振る舞いが、表示しているコンテンツの価値と関連しているかを分析するための実験を行った。

本実験で対象としたのは、Twitter から取得したコンテンツで、データセットを作成した。具体的には、2013 年 12 月 3 日にランダムで選んだ 56 個ツイートをデータセットと設定した。

被験者に、スマートフォンでデータセットの個々のコンテンツを読んでもらい、閲覧後に被験者に各コンテンツに関する価値を 1 点（最も価値が低い）から 5 点（最も価値が高い）までを評価してもらった。被験者は学部、大学院に所属し、常に twitter を利用する日本学生 3 名である。

ユーザの振る舞いのパラメータとして、フリック速度を利用する。フリック速度とは、スマートフォンでフリック操作をする際の指の移動速度のことであり、

具体的にはコンテンツ*i*に対するフリック速度 fv_i を以下の式で定義する．

$$fv_i = \frac{dist_i}{dur_i} \dots (2)$$

ここで， $dist_i$ は，コンテンツ*i*から次のコンテンツを表示するためのフリック操作の際の指の移動距離（ピクセル）であり， dur_i はその移動に要した時間（ミリ秒）である．

個々のコンテンツを閲覧した際のフリック速度を測り，フリック速度とコンテンツの価値の関係性を持っているかを分析する．各被験者によるコンテンツの振る舞い（フリック速度）とそのコンテンツの評価を獲得し，コンテンツの価値の評価によって各被験者のフリック速度の平均を図5に示す．

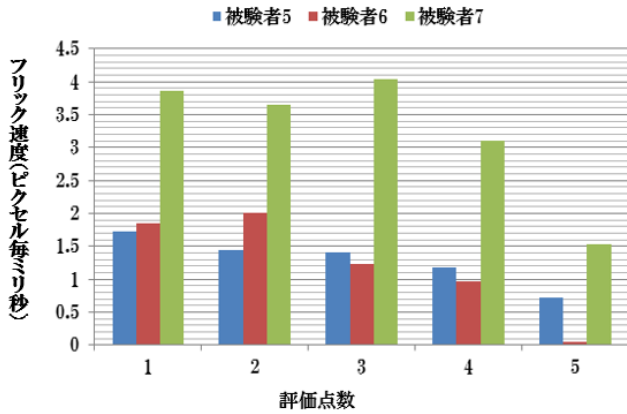


図 5. 各被験者の平均フリック速度

4.4 考察

上記の実験結果に対する考察を以下に示す．

- 価値が高いコンテンツの場合，読み速度が遅くなる傾向となる．逆に，価値が低いコンテンツの場合，読み速度が速くなる傾向となる．
- 価値が高いコンテンツほど，フリック速度が遅くなる傾向がある．逆に，価値が低いコンテンツの場合，フリック速度が速くなる傾向がある．
- ユーザや，ユーザの使用言語とは関係なく，コンテンツに対する振る舞いはコンテンツの価値を反映させると考えられる．

ユーザの振る舞いからを獲得することによりコンテンツの価値を推測作成し推薦する方法と，直接入力法の差異は少なく，提案手法によって，ユーザの振る舞いにより，ユーザへの負担が少ないやり方で，コン

テンツの価値を効果的に推測できることが期待できることが示せた．

5. 提案手法

本手法では，対象問題は，対象ユーザuが，読んでいないコンテンツ(Un-Visited Contents)の価値 V_{uc} を推測する．図6のように，ユーザ間の類似度を計算するために，まず対象ユーザにとって他のユーザの貢献度を計算することが必要である．そして，候補者と対象ユーザの振る舞いをマッチングすることにより，総合的な類似度を求める．最後に，未読コンテンツの価値を推定する．

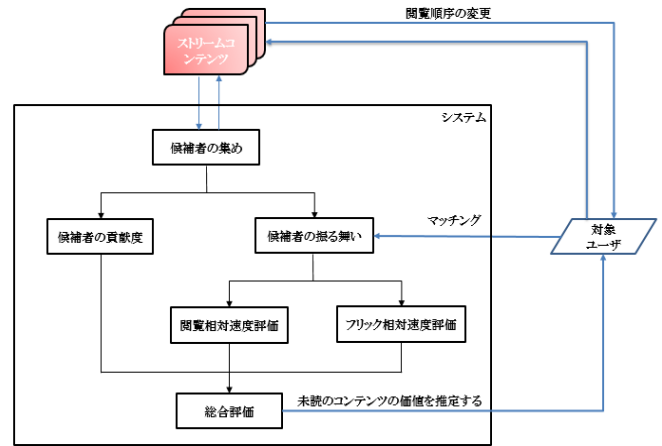


図 6. 類似度の計算の全体図

ここで，ストリームコンテンツ $C = \{c_1, c_2, \dots, c_n\}$ に対して，閲覧際の振る舞い $B = \{b_1, b_2, \dots, b_n\}$ と評価値 $E = \{e_1, e_2, \dots, e_n\}$ を獲得した．ユーザuは $u = (C_u, B_u, E_u)$ として表現する．なお，振る舞い**b**は， $b = (v_r, v_f)$ として，読み速度とフリック速度の組で表現する． v_r, v_f, e は以下のように正規化する．

$$v_{ri} = \frac{v_{ori} - \bar{v}_r}{\sigma} \dots (3)$$

$$v_{fi} = \frac{v_{ofi} - \bar{v}_f}{\sigma} \dots (4)$$

$$e_i = \frac{e_{oi} - \bar{e}}{\sigma} \dots (5)$$

ここで， v_{ori} は v_{ri} の正規化前の値であり， $\bar{v}_r$ は平均値を表し， σ は標準偏差を表す．(4), (5)も同様である．

5.1 候補者の貢献度

対象ユーザが閲覧したコンテンツに対して、それを読んでいないユーザを集めて候補者とし、その中から対象ユーザにとって貢献度が高いユーザを選別する。選別の際には、以下の仮説を利用する。

- 対象ユーザと候補者が共に閲覧したコンテンツの数が多いほど、その候補者は対象ユーザにとって有益である。
- 対象ユーザと候補者がどちらか一方しか閲覧していないコンテンツの数が少ないほど有益である。
- 受信するコンテンツが少ない候補者は、貢献度が低く、ノイズとして削除すべきである。

上記の考えに基づいて、対象ユーザ u にとって候補者 u_k の貢献度を以下の式で計算する。

$$Contribution(u, u_k) = \frac{|C_u \cap C_{u_k}|}{|C_u \cup C_{u_k}|} \dots (6)$$

ここで、 C_u と C_{u_k} は u と u_k が閲覧したコンテンツ(Visited Contents)を表す。

5.2 類似度の計算

コサイン類似度を用いて、対象ユーザと候補者の類似度を計算する。コサイン類似度は、ベクトル空間において、ユーザ同士を比較する際に用いられる類似度計算手法の一つであり、以下の式で定義される。

$$sim(u, u_k) = \cos\theta = \frac{\vec{u} \cdot \vec{u_k}}{|\vec{u}| \cdot |\vec{u_k}|} \dots (7)$$

図7のように、 u と u_k の類似度を計算し、1に近けれ

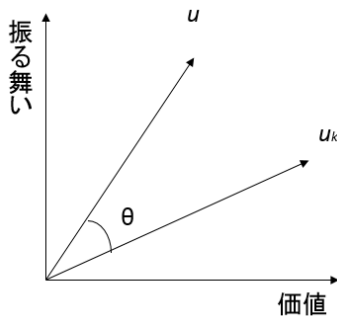


図7.類似度のイメージ

ば類似しており、0に近ければ似ていないことになる。

5.3 未読コンテンツの価値の推測

コンテンツに関する振る舞いを利用して、個々のコンテンツの価値の尺度とし、候補者が閲覧したそれぞれのコンテンツの振る舞いとその評価に基づいて推測する。

それを解決するために、まず振る舞いに対応する評価値に基づいて *fitting curve*^[30]を求める。獲得した振る舞いとその対応する評価値 $(v_{ri}, e_i)(i=1,2,\dots,n)$ がある。

$$P_m(x) = a_0 + a_1x + \dots + a_mx^m \dots (8)$$

関数 $e_r' = fc(v_r)$ をフィッティングする。以下の関数値が最小値になるような a_i を求める。

$$F(a_0, a_1, \dots, a_m) = \sum_{i=1}^n \sigma^2 = \sum_{i=1}^n [e_i - P_m(v_{ri})]^2 \dots (9)$$

そして

$$\begin{aligned} \frac{\partial F}{\partial a_j} &= 2 \sum_{i=1}^n [e_i - P_m(v_{ri})] v_{ri}^j \\ &= 2 \sum_{i=0}^n \left[e_i - \sum_{k=0}^m a_k v_{ri}^k \right] v_{ri}^j = 0 \dots (10) \end{aligned}$$

であることから、以下の条件が求まる。

$$\sum_{k=0}^m a_k \left(\sum_{i=1}^n v_{ri}^{k+j} \right) = \sum_{i=1}^n e_i v_{ri}^j \dots (11)$$

同様に、 $e_f' = fc(v_f)$ の近似曲線も求める。候補者が自分の振る舞いから自分に対して総合的に価値 e' を推定する。

$$e' = \alpha \cdot e_r + (1 - \alpha) \cdot e_f \quad \alpha \in [0,1] \dots (12)$$

最後に、ユーザ同士の類似度に基づいて対象ユーザに対して未読コンテンツの価値 V_{uc} を推測する。

$$V_{uc} = \bar{e}_u + \frac{\sum_{k=1}^N \text{sim}(u, u_k) * (e'_{uc} - \bar{e}_{u_k})}{\sum_{k=1}^N (\text{sim}(u, u_k))} \dots (13)$$

$\bar{e}_u$ は獲得した対象ユーザのすべての評価の平均、 $\bar{e}_{u_k}$ は獲得した候補者のすべての評価の平均、 e'_{uc} は候補者 u_k の未読コンテンツの価値の推定値である。

6. 評価実験

ユーザの振る舞い協調フィルタリングによるソーシャルストリームにコンテンツの価値の推測方法を評価するため、提案手法と *TF-IDF*^[31]法と *TF-IDF*を用いる協調フィルタリング手法とによる推測の評価ごとの平均絶対誤差 *MAE*^[32]で有効性を検証する。対抗手法は、ユーザが閲覧した内容に基づいて、*TF-IDF*のテキスト処理を用いて、ユーザの嗜好との類似を計算し、未読のソーシャルストリームコンテンツの価値を推測する。ところで、対抗手法では対象ユーザの類似度を以下の式で計算する。

$$\text{sim}(u, u_k) = \frac{\sum_{i=1}^n |e_{u_i} - e_{u_{k_i}}|}{n} \dots (14)$$

そこで、平均絶対誤差 *MAE* (Mean Absolute Error) は数値予測問題における推測する質を評価する指標の1つである。予測値が実際の評価値から平均的にどの程度の乖離を表す。

$$\text{MAE} = \frac{\sum_{i=1}^n |e_i - e'_i| \pm \varepsilon}{n} \quad \varepsilon \in [0, 0.1] \dots (15)$$

ε は誤差となる。実験の参加者は10人、各被験者が閲覧するコンテンツが100個となった。各手法による *MAE* の結果は図7に示す。なお、提案手法においてソーシャルストリームコンテンツの数量によって増えるごとに推測する質が高くなった。また、類似度が高い候補者が増えるごとに平均絶対誤差が減少していることがわかる。しかし、データ量の増加による算出する量も増加し、異なるユーザによって振る舞いの類似性が一部の偏りを算出されるであろう。次のステップの問題として検討することが必要であるが、現段階では、

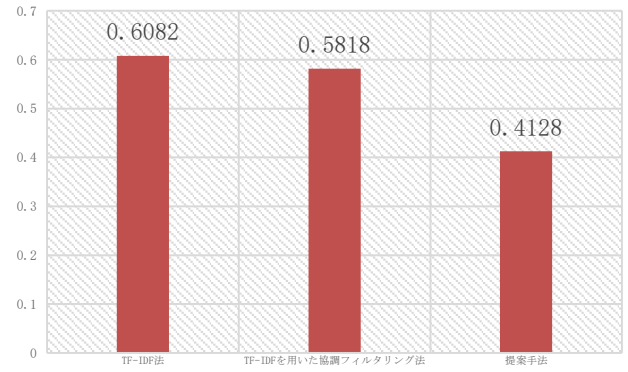


図7. 対抗手法と提案手法の比較

振る舞い協調フィルタリングによるコンテンツの価値の推測は現実的であり、効果が比較的理想的である。

7. まとめ

本稿では、スマートフォンでのSNSにおいて、ユーザの振る舞いをコンテンツの価値・重要度の度合いとして測定し、対象ユーザの未読コンテンツに対する適切な提示時間・順序を推定することにより、効率な閲覧の可能な手法について述べた。この手法では、あらかじめストリームコンテンツの中からユーザ同士の振る舞いが類似しているかを推定することにより、対象ユーザの閲覧の支援をする。このアプローチを実現するためには、ユーザの振る舞いにもとづいてコンテンツの価値を推定することが必要である。本論文では、コンテンツ閲覧時の、読む速度とフリック速度とコンテンツの価値の間に相関がある可能性があることを被験者実験により示した。

今後の展望として、ソーシャルストリームコンテンツの価値を推測した後、より効率的な提示する手法を開発する必要がある。また、本研究の対象は、SNSによるソーシャルストリームのため、リアルタイムに推定する手法が必要だと考えられる。メモリベースの協調フィルタリングの特徴は、推測の質が低い、処理速度が速い。一方、モデルベースの協調フィルタリングの特徴は、推測の質が高い、処理速度が遅い。そのため、ハイブリッド型の協調フィルタリング^[6]を用いることで、効果的な推測の質や処理速度を実現することが可能と考えられ、効率的な閲覧が更なる向上をもたらすことが期待できる。

参考文献

- [1] P. Arabie, and Y. Wind, "Marketing and Social Networks." In Stanley Wasserman and Joseph Galaskiewicz, *Advances in Social Network Analysis:*

Research in the Social and Behavioral Sciences. Thousand Oaks, Calif.: Sage Publications, pp. 254–273, 1994.

- [2] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of Twitter data," *Proceedings of the Workshop on Languages in Social Media*, 2011:30-38.
- [3] N. J. Belkin, and W. B. Croft, "Information filtering and information retrieval: two sides of the same Coin," *Communication of the ACM* 1992, 2(1):134-145.
- [4] M. Balabanović, "An adaptive web page recommendation service," *Proceedings of the first International Conference on Autonomous Agents*. 1997: 378-385.
- [5] C. Basu, H. Hirsh, and W. Cohen, "Recommendation as classification: using social and content based information in recommendation," *Proceedings of the 1998 workshop on Recommender Systems*, 1998: 11-15.
- [6] D. R. Liu, and Y. Y. Shih, "Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences," *Journal of Systems and Software*, vol.77, no.2, pp:181-191, 2005.
- [7] P. Resnick, and H. R. Varian, "Recommender Systems," *Communication of the ACM*, vol.40, no.3, pp:56-58, 1997.
- [8] J. J. Rocchio, "Relevance Feedback in Information Retrieval," *the Smart Retrieval System: Experiments in Automated Document Processing*, Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1971.
- [9] S. Park, and D. M. Pennock, "Applying collaborative filtering techniques to movie search for better ranking and browsing," *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2007: 550-559 .
- [10] J. Delgado, and N. Ishii, "Memory-Based Weighted Majority Prediction for Recommender Systems," *Proceedings of the ACM SIGIR 1999 Workshop Recommender Systems: Algorithms And Evaluation*, 1999.
- [11] R. F. Jia, and M. Z. Jin, and C. Liu, "A new clustering method for collaborative filtering," *Proceedings of International Conference on Networking and Information Technology*, 2010: 488-492.
- [12] S. J. Gong, and H. W. Ye, "Joining user clustering and item based collaborative filtering in personalized recommendation services," *Proceedings of International Conference on Industrial and Information Systems*, 2009: 149-151.
- [13] H. Lee, Y. S. Choi, S. Lee, and E. Shim, "Smart pose: mobile posture-aware system for lowering physical health risk of smartphone users," *Proceedings of CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2013:2257-2266.
- [14] 土岐真里奈, 牛尼剛聡, "ソーシャルストリーム閲覧時の振舞いを利用したユーザプロファイル構成手法," *情報処理学会論文誌データベース*, vol.6, no.4, pp: 35-45, 2013.
- [15] S. Amini, Y. Li, "Crowd Learner: rapidly creating mobile recognizers using crowdsourcing," *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013: 163-172.
- [16] Q. Xu, J. Erman, A. Gerber, Z. Mao, J. Pang, and S. Venkataraman, "Identifying diverse usage behaviors of smartphone apps," *Proceedings of the 2011 ACM SIGCOMM conference on Internet Measurement Conference*, 2011: 329-344.
- [17] 甲谷優, 岩田具治, 藤村考, "複数人によるアカウントの共有を考慮したトピックモデルに基づく協調フィルタリング," *DBSJ Journal*, vol.9, no.3, pp:7-12, 2011.
- [18] J. S. Breese, D. Heckerman, and C. Kaide, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998 : 43-52
- [19] L. H. Ungar, and D. P. Foster, "Clustering Methods for Collaborative Filtering," *Workshop on Recommender Systems at the 15th National Conference on Artificial Intelligence*, 1998 : 112-125.
- [20] M. Connor, and J. Herlocker, "Clustering Items for Collaborative Filtering," *Proceedings of SIGIR 2001 Workshop on Recommender System*, 2001.
- [21] J. B. Schafer, J. Konstan, and J. Riedl, "Recommender systems in E-commerce," *Proceedings of the first ACM Conference on Electronic Commerce*, 1999:158-166.
- [22] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of Recommendation Algorithms for E-Commerce," *Proceedings of the 2nd ACM Conference on Electronic Commerce*, 2000 : 158-167.
- [23] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," *Proceedings of the 10th International Conference on World Wide Web*, 2001:285-295
- [24] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, 1999: 230-237.
- [25] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e-commerce," *Proceedings of the 2nd ACM conference on Electronic commerce*, 2000:158-167.
- [26] J. Herlocker, J. A. Konstan, and J. Riedl, "An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms," *Information Retrieval*, 2002: 287-310.
- [27] A. Singhal, "Modern Information Retrieval: A Brief Overview". *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 24 (4): 35–43.
- [28] P. N. Tan, M. Steinbach and V. Kumar, "Introduction to Data Mining," Addison-Wesley.
- [29] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," in *The American Statistician*, 42(1):59–66, February 1988.
- [30] A. Ralston and H. S. Wilf, *Mathematical Methods for Digital Computers* John Wiley & Sons, New York, 1960.
- [31] G. Salton and M. J. McGill: *Introduction to modern information retrieval*. McGraw-Hill, ISBN 0-07-054484-0.
- [32] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp: 679-688, 2006.