

フォロー関係を考慮したゲーム推薦

田村麻衣¹ 清雄一¹ 田原康之¹ 大須賀昭彦¹

概要：本稿では、数多くの無料ゲームアプリケーションソフトウェアをデータとして、従来の協調フィルタリングとは異なる手法で情報推薦を行う。従来の方法では評価数が少ないアイテムをデータセットに用いる場合、その精度は低い。そこで、各ユーザが実際に遊んでいるゲーム、そしてユーザ間のフォロー関係を考慮して協調フィルタリングを行う。各ゲームにおいて遊んでいる・いないを疑似的な評価値とし類似度を求め、そこにフォロー関係の有無による補正値を加え、フォロー関係を考慮しなかった場合と比較する。その結果、フォロー関係を考慮した方が、ユーザが元々登録しているゲームがより高い順位に位置することがわかった。

キーワード：情報処理学会論文誌ジャーナル、情報推薦、協調フィルタリング、ゲーム

1. はじめに

スマートフォンの普及に伴い、スマートフォン向けのアプリケーションソフトウェア、通称「アプリ」の開発が急速に進んでいる。AppleのApp Storeでは過去10年間に1700億回のダウンロードがあり、消費者が使った金額は1300億ドルを超えている[1]。中でも無料のゲームアプリの人気は一部社会現象を引き起こすまでになっている。数多くの無料ゲームアプリが開発されるにつれ、攻略情報を掲載するWebサイトやゲームのレビューを書いたブログ、おすすめのゲームを紹介するWebサイトなどが登場してきた。

そんな多くのコンテンツの中から、ユーザ自身の嗜好に沿ったものを見つける方法として情報推薦システムの利用が期待されている。情報推薦システムはリコメンダシステム(Recommender System)あるいはリコメンデーションシステム(Recommendation System)とも呼ばれ、CACM[2]で紹介されたことで広く知られるようになった[3]。情報推薦システムの典型的な手法として、協調フィルタリング法(Collaborative Filtering Method)がある[4]。ユーザベース協調フィルタリングは、ユーザベースの推薦手法の中で最も多用される手法の一つであるが[5]、推薦の精度と多様性の両立が難しいとされており[6]、多くの研究者が推薦制度と多様性の両立を目的とした研究を行っている[7-11]。一般的にユーザベースの推薦手法では、ユーザがアイテムに与えた評価値を基に評価指標を用いてユーザ間の類似度を求めることで情報推薦を行っている。しかし、その精度はアイテムへの評価数に影響されるため、評価数が少ないアイテム

を扱う場合適切な推薦を行うことは困難である。

そこで本研究では、アイテムに対する評価値以外の指標を用いて情報推薦を行うことを目的とする。

既存の検索エンジンで「ゲームアプリ サイト」と検索を行うと、「GameWith-国内最大級のゲームアプリ情報・攻略サイト」と書かれたページが一番上に表示される[12]。これには、ゲームの概略や攻略情報の他、会員登録しているユーザが書いたゲームへのレビューや評価などが掲載されている。GameWithの特徴として、会員登録しているユーザ自身が遊んでいるゲームを登録することで、そのゲームのコミュニティに入ることができる。また、他のユーザをフォローすることができ、フォローしたユーザが投稿したフィードが簡単に閲覧できるようになっている。

本研究では、各ゲームへの評価値ではなく、ユーザが登録しているゲームとユーザ間のフォロー関係を指標とした情報推薦を行うことを目指す。

2. 関連研究

2.1 ゲーム推薦

ゲームの多様化によってユーザに合わせたゲームを推薦する推薦システムの研究が行われている[13-14]。中谷ら[13]は、Amazon.comなどのECサイトに掲載されているレビュー文を解析することでユーザの「経験値」を取得し、それを用いてゲーム推薦を行っている。経験的価値を感覚的知覚、創造的思考、身体性、社会性、衝動的感情の5つに分類し、作成した辞書を用いてテキストマイニングした結果、従来

¹ 電気通信大学
The University of Electro-Communications, Chofu,
Tokyo 182-8585, Japan

の推薦法とは大きく異なる結果が示されたと報告している。また、大山ら[14]はゲームレビューサイトに掲載されているレビュー文を Word2vec に学習させ、言葉の足し引きによってゲームを定量化し、ゲームを推薦するシステムの研究を行っている。

これらのシステムはレビュー文の量によって精度に差が生まれ、レビューを書いていないユーザやレビューが書かれていないゲームを対象に推薦を行うことは困難である。

2.2 協調フィルタリング

本節では協調フィルタリングについての概略を述べる。本節の内容は文献[4]に基づく。

協調フィルタリングによる嗜好の予測を行う。ここでいう予測とは、ユーザがまだ知らない、他のユーザは知っているアイテムについて、ユーザの関心の有無や評価値を推定することである[4]。この予測は主に次の2段階で実現する。

2.2.1 類似度の計算

n 人の全ユーザの集合を $X = \{1, \dots, n\}$ 、 m 種類の全アイテムの集合をとす。評価値行列 R はユーザ $x \in X$ の、アイテム $y \in Y$ への評価値 r_{xy} を要素とする行列である。 r_{xy} は評価済みなら評価値の定義域 R のいずれかの値をとり、未評価なら欠損値「*」をとす。 $R = \{1, 2, 3\}$ の評価値行列の例を表1に示す。

表1 評価値行列の例

ユーザ	アイテム			
	A	B	C	D
ユーザ1	1	3	*	3
ユーザ2	*	1	3	*
ユーザ3	2	1	3	1
ユーザ4	1	3	2	*

ユーザを添え字 a で表す、すなわち、 r_{ay} はユーザのアイテム y への評価値である。また、ユーザ x が評価済みのアイテムの集合を $Y_x = \{y \mid y \in Y, r_{xy} \neq *\}$ と表す。ユーザ a とユーザ x の類似度は、両者が共に評価しているアイテムについての Pearson 関数で測る。

$$\rho_{ax} = \frac{\sum_{y \in Y_{ax}} (r_{ay} - \bar{r}_a')(r_{xy} - \bar{r}_x')}{\sqrt{\sum_{y \in Y_{ax}} (r_{ay} - \bar{r}_a')^2} \sqrt{\sum_{y \in Y_{ax}} (r_{xy} - \bar{r}_x')^2}} \quad (3.1)$$

ただし、 Y_{ax} はユーザ a とユーザ x が共通に評価したアイテムの集合、すなわち $Y_{ax} = Y_a \cap Y_x$ で、 $\bar{r}_x' = \sum_{y \in Y_{ax}} r_{xy} / |Y_{ax}|$ である。なお、 $|Y_{ax}| \leq 1$ 、すなわち、ユーザ a とユーザ x が共通に

評価したアイテムが一つ以下ならば、Pearson 関数は計算できないため、 $\rho_{ax} = 0$ とする。

2.2.2 嗜好の予測

アイテム $y \in Y_a$ の評価値は式(3.1)の類似度で重みづけした、各ユーザのアイテム y への評価値の加重平均で予測する。

$$\hat{r}_{ay} = \bar{r}_a + \frac{\sum_{x \in X_y} \rho_{ax} (r_{xy} - \bar{r}_x)}{\sum_{x \in X_y} |\rho_{ax}|} \quad (3.2)$$

ただし、 X_y はアイテム y を評価済みの利用者の集合で \bar{r}_x はユーザ x の全評価アイテムに対する平均評価値 $\bar{r}_x = \sum_{y \in Y_x} r_{xy} / |Y_x|$ である。なお、ユーザがアイテム y を評価済みである $a \in X_y$ である状況ではそもそも r_{ay} の推定が不要になるので想定しなくてよい。この式の第1項は、第2項が中間的な評価で0をとるので、それを補正するバイアス項である。また、第2項の分子は上記の加重平均であり、分母は評価しているユーザが多い、すなわち $|X_y|$ が大きいと加重平均が大きくなりやすい問題に対する正規化項である。

以上の手順を用いて対象ユーザの未評価アイテムそれぞれに対して予測を行い、高い予測値を持つアイテムを推薦することができる。

2.3 SNS やブログのフォロー形成

近年 Twitter や Facebook などの SNS を利用するユーザの数は急速に伸びており、著名人や企業による宣伝活動、災害時の情報伝達の手段として利用されるなど、大きな注目を浴びている[15]。Twitter の特徴として、興味がある内容を投稿しているユーザをフォローするだけで、投稿者のツイートを閲覧することができる。ユーザがフォロー関係を形成する理由に関する研究が行われており、小出ら[15]は、フォロー関係特有のネットワークを対象に、その中心的な役割を果たしているユーザに着目してユーザ間のフォロー、被フォローの類似関係における特徴を、ブログやレビューサイトにおける読み手、書き手によるネットワークと比較することで分析を行っている。Twitter、日本のブログサービスサイトである“Ameba” (以下、ameblo)、化粧品の口コミサイト“@cosme” (以下、cosme) の三つにおけるそれぞれのネットワークの相互性を比較したところ、Twitter は 60%、ameblo は 54%、cosme は 44% と、口コミサイトがやや低い値となったと報告している。

3. 提案手法

本節では協調フィルタリングの考え方をういた上で、ユーザ間のフォロー関係を考慮した推薦手法を提案する。対象ユーザの選好パターンに基づいてユーザ間の類似度を求める際、フォロー関係を考慮することで、評価数が少ないアイテムを扱う場合でも高い精度で情報推薦を行う。

3.1 提案手法の概要

2.1 節の協調フィルタリングの考え方を適用し、以下の手順で推薦を行う。提案手法の概要図を図 1 に示す。

まず、対象ユーザが登録しているゲームを 1、登録していないゲームを 0 とし、疑似的に評価値を定め評価値行列を作成する。式(3.1)を用いてユーザ a とユーザ x の類似度 ρ_{ax} を求める。また、ユーザ a とユーザ x をフォローしていた場合、類似度に補正值 p_{ax} を加え、補正後の値 ρ'_{ax} を用いて推定値を算出する。

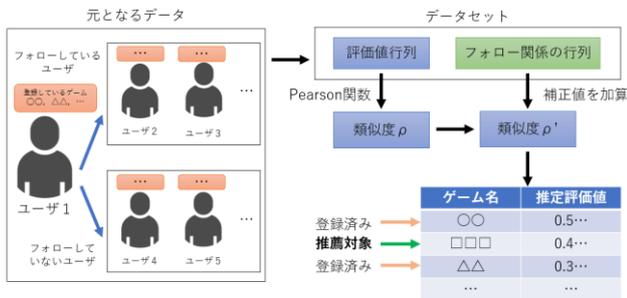


図 1 提案手法の概要図

3.2 補正值 p_{ax} の算出

例として、本節では、補正值 p_{ax} の求め方について述べる。まず、 n 人の各ユーザのフォロー関係を表す $n \times n$ 行列を作成するため、フォロー関係におけるユーザ間の距離 f を求める。フォロー関係についての概要図を図 2 に示す。

ユーザ a がユーザ x をフォローしていた場合、 $f_{ax} = 1$ とする。また、フォローしているユーザがフォローしているユーザ、といったように、数段階でのフォロー関係を考慮するため、一段階目の関係（直接フォローしているユーザ）は $f = 1$ 、二段階目の関係（フォローしているユーザがフォローしているユーザ）は $f = 0.5$ 、以降一段階こえるたびに半減させる。なお、自分自身への距離、およびフォロー関係がない場合は $f = 0$ とする。

3.3 ユーザ間の距離の算出

本節では、フォロー関係におけるユーザ間の距離 f を求める手順を示す。

n 人のフォロー関係を表す $n \times n$ 行列 M を作成する。ユーザ i とユーザ j の1段階目のフォロー関係を $m_{i,j}^{(1)}$ と表すとき、

$$M = \begin{bmatrix} m_{1,1}^{(1)} & \dots & m_{1,n}^{(1)} \\ \vdots & \ddots & \vdots \\ m_{n,1}^{(1)} & \dots & m_{n,n}^{(1)} \end{bmatrix} = F^{(1)}$$

となる。なお、 $m_{i,j}^{(1)} = f_{i,j}^{(1)}$ である。2段階目のフォロー関係を考えると、行列 M の2乗は

$$M^2 = \begin{bmatrix} m_{1,1}^{(2)} & \dots & m_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ m_{n,1}^{(2)} & \dots & m_{n,n}^{(2)} \end{bmatrix}$$

であり、このとき2段階目までフォロー関係を考慮した行列 $F^{(2)}$ は

$$F^{(2)} = \begin{bmatrix} f_{1,1}^{(2)} & \dots & f_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ f_{n,1}^{(2)} & \dots & f_{n,n}^{(2)} \end{bmatrix}$$

で表され、このとき $f_{i,j}^{(2)}$ は $f_{i,j}^{(1)}$ が0以上ならば $f_{i,j}^{(1)}$ の値をとり、 $f_{i,j}^{(1)}$ が0で $m_{i,j}^{(2)}$ が0ならば0、 $f_{i,j}^{(1)}$ が0で $m_{i,j}^{(2)}$ が0以上ならば 2^{-1} となる。

同様に、 t 段階目のフォロー関係を考えると、行列 M の t 乗は

$$M^t = \begin{bmatrix} m_{1,1}^{(t)} & \dots & m_{1,n}^{(t)} \\ \vdots & \ddots & \vdots \\ m_{n,1}^{(t)} & \dots & m_{n,n}^{(t)} \end{bmatrix}$$

となり、 t 段階目までフォロー関係を考慮した行列 $F^{(t)}$ は

$$F^{(t)} = \begin{bmatrix} f_{1,1}^{(t)} & \dots & f_{1,n}^{(t)} \\ \vdots & \ddots & \vdots \\ f_{n,1}^{(t)} & \dots & f_{n,n}^{(t)} \end{bmatrix}$$

である。このとき $f_{i,j}^{(t)}$ は $f_{i,j}^{(t-1)}$ が0以上ならば $f_{i,j}^{(t-1)}$ の値をとり、 $f_{i,j}^{(t-1)}$ が0で $m_{i,j}^{(t)}$ が0ならば0、 $f_{i,j}^{(t-1)}$ が0で $m_{i,j}^{(t)}$ が0以上ならば $2^{-(t-1)}$ となる。

例として5人のユーザのフォロー関係が表2のようになっていた場合、複数段階フォロー関係を考慮すると表3のようになる。

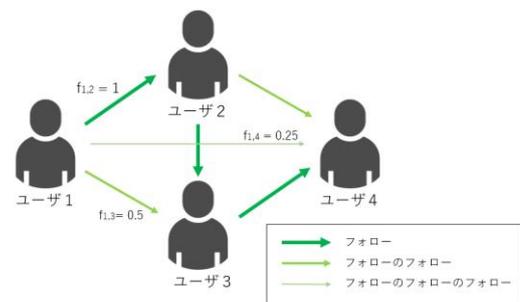


図 2 フォロー関係を数段階考慮する場合

表 2 5人のユーザのフォロー関係

	ユーザ1	ユーザ2	ユーザ3	ユーザ4	ユーザ5
ユーザ1		1			
ユーザ2			1		1
ユーザ3				1	1
ユーザ4		1			
ユーザ5		1		1	

表 3 複数段階フォロー関係を考慮した場合

	ユーザ1	ユーザ2	ユーザ3	ユーザ4	ユーザ5
ユーザ1	0	1	0.5	0.25	0.5
ユーザ2	0	0	1	0.5	1
ユーザ3	0	0.5	0	1	1
ユーザ4	0	1	0.5	0	0.5
ユーザ5	0	0.5	0.25	1	0

この行列をもとに補正值 p_{ax} を各ユーザ間の類似度に加算する。先行研究によると SNS 等に比べ口コミサイトにおけるフォロー関係の相互性はやや低い。同じコミュニティに所属しているなど関連性のあるユーザ同士でフォローネットワークが形成されているわけではなく、各ユーザが個人的に興味を持ったユーザをフォローしていると考えられる。すなわち、フォロー関係と登録ゲームの相関は低い。そこで本稿では、フォローしているユーザとの類似度に負の補正值を加えることで、より高い精度で情報推薦を行うことができるかと仮定し、実験を行う。

4. 実験

提案手法の有用性を検証するために実験を行った。本章では実験に使用したデータセットの詳細と統計量、また実験の内容と評価について述べる。

4.1 データセット

本実験では、GameWith にアカウントを新規登録し、「おすすめユーザ」として表示されたユーザの中から、3つ以上のゲームを登録しているユーザを 100 人選び、データセットを作成した。なお、アイテムは選択したユーザが一人以上登録しているゲーム(全 107 種)のみとした。このデータセットにおける統計量を表 4 にまとめた。

表 4 データセットにおける統計量

項目名	平均	最大値	最小値
登録しているゲーム数	5.91	17	3
フォローしているユーザの数	3.31	23	0

また、登録者数が最も多いゲームは“モンスターストライク”

で 100 人中 57 人、次いで“パズル&ドラゴンズ”が 39 人、“バンドリ！”が 33 人であった。

4.2 実験結果

式(3.1)を用いて全ユーザ間の類似度を求め、これに補正值を加えた場合と加えなかった場合の、2種類の類似度 A, B を用意した。補正值に関して、四段階までフォロー関係を考慮して作成したフォロー関係を表す行列に、-0.5 を掛けたものを各類似度に加算した。類似度 A, B について、各ユーザの全 107 種のアイテムに対する推定評価値を求め、降順に並び替えたとき、上位 5 個、10 個、20 個、30 個のゲームの中に元々各ユーザが登録していたゲームが含まれる個数(以下、該当数)をそれぞれ比較した。類似度 A を用いた場合に、類似度 B を用いた場合と比べたとき、該当数が増えた人数と減った人数をまとめた。

4 か所で区切り比較したが、どれを見ても類似度 A のほうが類似度 B に比べて該当数が増えていることがわかった。なお、ほとんどの場合で増減したゲームの個数は 1 つであったが、上位 5 個で比較したとき該当数が 2 つ増えたユーザが 2 人いた。

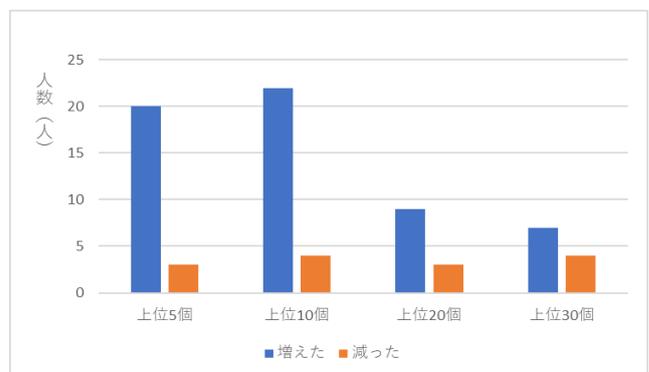


図 3 該当数の増減

5. 考察

実験結果によると、フォローしているユーザとの類似度を低くすることで、元々各ユーザが登録していたゲームが上位に位置するようになることがわかった。つまり、ユーザが登録しているゲームとフォローしているユーザが登録しているゲームが異なっている事が多いということがわかる。自分が遊んでいなくとも興味があるゲームに関する投稿を行っているユーザなど、直接的な関連性が低いユーザがフォローしているユーザに含まれると考えられる。

また、該当数が減ったユーザに関して、類似度 B を使用した場合には上位 5 個、10 個に含まれていて、類似度 A では含まれなかったゲームを見てみると、“モンスターストライク”、“バンドリ!”、“ラブライブ”、“ゼノンソード”の 4 つであった。前 2 つは登録者数が 1 位と 3 位であり、フォ

ロー関係の有無に関わらず登録している可能性が高いため、類似度を減らしたことで除外されてしまったと考えられる。このように、全体として人気のゲームの場合、フォロー関係との関連性が低く、フォロー関係を考慮して類似度を減らした上での推定評価値から推薦する以外の方法を用いる必要がある。例えば、純粋に登録者数が多いものとして推薦するなどといった方法がある。後2つに関して、GameWithにおけるコミュニティの人数はどちらも200人前後であった。“モンスターストライク”の約20万人、“バンドリ!”の1万5千人と比較すると、その規模はやや小規模である。人気のゲームに比べて出回っている情報量が少ないことを考えると、小規模なコミュニティを形成しているゲームを遊んでいるユーザー同士は相互的なフォロー関係を築いており、登録ゲームとフォロー関係の相関が高いのではないかと考えられる。

6. おわりに

本稿では、従来のユーザーのアイテムに対する評価値を用いる従来の手法では対応できない、評価数が少ないゲームをデータにした場合に情報推薦を行う手法を提案した。その結果、フォローしていることと同じゲームを遊んでいることの関連性は低いということがわかった。

今後は、データ数を増やすとともに、フォロー関係の詳細な分析を行い、より精度の高い手法の構築を目指す。また、フォロー関係以外にもデータとして用いることができる要素がないかを検討する。

謝辞 本研究は JSPS 科研費 JP17H04705, JP18H03229, JP18H03340, JP18K19835, JP19H04113, JP19K12107 の助成を受けたものです。

本研究を遂行するにあたり、データ収集にご協力いただいた GameWith を運営する方々、及び活発な議論と貴重な御意見を頂いた研究グループの皆様へ感謝致します。

参考文献

[1] “iOS App Store、2010年以降のダウンロード数は1700億回、売り上げは1300億ドル”。
<https://jp.techcrunch.com/2018/06/01/2018-05-31-ios-app-store-has-seen-over-170b-downloads-over-130b-in-revenue-since-july-2010/>, (参照 2020-01-14)

[2] P. Resnick, and H. Varian,.: Recommender Systems, Comm. Of the ACM, Vol.40, No.3, pp-56-58 (1997).

[3] 樽井 勇之: 協調フィルタリングとコンテンツ分析を利用した観光地推薦手法の検討.

[4] 神尾 敏弘: 推薦システムのアルゴリズム..

[5] 山脇 淳一, 工藤 康生, 村井 哲也: 関係性マイニングと協調フィルタリングを用いた情報推薦手法. 日本感性工学会論文誌, Vol.17, No.4, pp481-488 (2018).

[6] Bradley, K., and Smyth, B.: Improving recommendation diversity, Proc. 12th Nat. Cogn. Sci. (AICS-01), Maynooth, Ireland, pp.75-84, 2001.

[7] Vozalis, M. G., Markos, A. I., and Margaritis, K, G.: A Hybrid Approach for Improving Prediction Coverage of Collaborative Filtering, Proc. of AIAI 2009, pp491-498, Springer, 2009.

[8] Zhou, T., Kuscsik, Z., Liu, J.G., Medo, M., Wakeling, J, R., and Zhang, Y. C.: Solving the apparent diversity-accuracy dilemma of recommender systems, Proceedings of the National Academy of Sciences of the United States of America, 107(10), pp4511-4515, 2010.

[9] Gan, M. X., and Jiang, R.: Constructing a user similarity network to remove adverse influence of popular objects for personalized recommendation. Expert Systems with Application, 40, pp.4044-4053, 2013.

[10] Gan, M. X., and Jiang, R.: Improving accuracy and diversity of personalized recommendation through power law adjustments of user similarities, Decision Support Systems, 55, pp.811-821, 2013.

[11] Zhang, Z. P., Kudo, Y., and Murai, T.: Neighbor selection for user-based collaborative filtering using covering-based rough sets, Annals of Operations Research, 256(2), pp.359-374, 2017.

[12] “GameWith”: <https://gamewith.jp/>. (参照 2020-01-14)

[13] 中谷 知博, 星野 准一: 経験的価値の分類に基づくゲーム推薦システム. 情報処理学会, 研究報告, 2008-EC-011, Vol.2008, No.129, pp49-56 (2008)

[14] 大山 浩暉, 竹川 佳成, 平田 圭二: レビュー文を考慮したゲーム推薦システムの実現に向けた単語の類似度調整の取り組み. エンタテインメントコンピューティングシンポジウム(EC2017), (2007)

[15] 小出 明弘, 斉藤 和巳, 風間 一洋, 鳥海 不二夫: ネットワーク分析による Twitter ユーザーのフォロー形成に関する一考察. 情報処理学会論文誌, 数理モデル化と応用, Vol.6, No.2, pp164-173(Aug. 2013).