

このユーザは信頼に値するか？

- SNS における協調的な信頼推定モデル -

南 大智[†] 牛尼 剛聡^{††}

[†]九州大学大学院芸術工学府 〒815-8540 福岡県福岡市南区塩原 4-9-1

^{††}九州大学大学院芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: [†]minami@kyudai.jp, ^{††}ushiama@design.kyushu-u.ac.jp

あらまし Twitter や Facebook などに代表される SNS の爆発的に普及に伴い、見知らぬ人同士がインターネットを介してコミュニケーションをとることが一般的に行われるようになった。また、近年では Airbnb や Uber などのシェアエコノミーに基づいたサービスが人気を獲得しており、見知らぬ人と繋がりを持つ機会はより増加している。しかし、ユーザはこれらのサービスを用いる上で、見知らぬユーザが本当に信頼できるかを判定したいという要求がある。しかし、見知らぬ他人が信頼できるかを判定するのが困難である場合が少なからず存在する。また、ユーザが行った判定が必ずしも正しいとは限らない。本研究では、対象ユーザに関する既知の情報を用いて、未知のユーザに対する信頼を予測するモデルを提案する。今回は提案モデルに基づき、オンラインレビューサイトにおけるレビュワーへの信頼の予測を行い、得られた信頼情報を用いて協調フィルタリングに基づくアイテム推薦を行い、有効性を評価する。**キーワード** 信頼推定, レビュー, Doc2Vec, 情報推薦, 協調フィルタリング

1. はじめに

1.1 背景

近年、Twitter や Facebook などのソーシャルネットワークサービス (SNS) が一般的に利用されるようになった。ここでは、現実の知り合いやコミュニティと繋がるために利用されるだけでなく、見知らぬ他人とインターネットを介して関係を構築し、コミュニケーションをとることも一般的に行われている。また近年、Airbnb [1] や Uber [2], TaskRabbit [3] などのシェアリングサービスの普及により、見知らぬ人同士がインターネットを媒介して繋がるだけでなく、現実世界でも繋がりコミュニケーションをとるようになった。このように、インターネットの発展により、見知らぬ他者と繋がりを持つ機会がより増加している。

1.2 解決すべき問題

しかし、インターネット上の全てのユーザが信頼できるとは限らない。中には、悪意を持ったユーザや対象ユーザから見て信頼に値しないユーザも存在する。ユーザは見知らぬ他者との関わりが必要とされるサービスを利用する上で、見知らぬユーザが信頼できるユーザであるかを判断することは困難である。

1.3 アプローチ

本研究では、対象ユーザの既知のユーザに対する信頼度に基づいて、未知のユーザに対する信頼度を予測するモデルを提案する。本モデルでは、対象ユーザが既知であるユーザに対して信頼度を評価し、評価されたユーザと未知のユーザとの類似性に基づき、未知のユーザに対する信頼度を推定する。

本論文では、我々が提案する信頼度推定モデルに基づき、オンラインレビューサイトにおけるレビュワーに対する信頼度を

推定し、信頼情報を用いた協調フィルタリングによる推薦を行う。具体的には、まず対象ユーザが選択した書籍に関するページにおいて、その書籍に対して投稿されたレビューを表示する。対象ユーザはレビューを読み、個々のレビュワーに対して支持するかどうかについての評価を入力する。評価されたレビュワーの特徴とまだ評価されていないレビュワーの特徴との間で類似度を計算することで、対象ユーザが未知であるレビュワーに対する信頼度を推定する。レビュワーの特徴は、あらかじめ学習した Doc2Vec を用いて、投稿した文書をベクトル化として表現する。また、対象ユーザが閲覧したレビューがどれだけそのレビュワーの典型的な文書であるかを考慮することで、推定される信頼度の確信性を向上させる。そして、各レビュワーに対する信頼度をレビュワーの重みとして、推薦書籍の計算および提示を行う。

2. 関連研究

2.1 社会的な信頼推定

これまでに、レビューなどのテキスト情報の信頼性に関して、さまざまな研究が行われてきた [4-7]。これらの研究では、スパムなど、不当なレビューの投稿を行う悪意を持った投稿者を問題視しており、「スパムであるか否か」「虚偽の投稿であるか否か」ということに関する議論を展開している。また、ユーザの信頼性に関する研究として、小林ら [8] は、インターネットオークションにおけるユーザ間の関係性を考慮することで、ユーザの信頼性評価を行なった。上記の研究において議論されているのは、社会的な観点から見た信頼性である。一方で、「対象ユーザにとって有益なユーザであるか」や「対象ユーザが好感を抱



図1 インターフェース上での評価の入力例

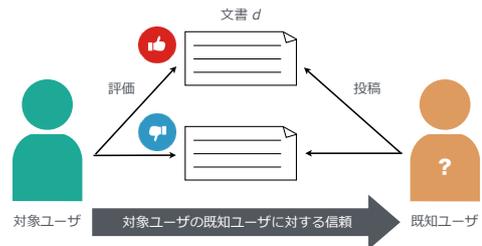


図2 既知のユーザに対する信頼

くか」など、個人的な観点から見た信頼性についてはあまり議論されていない。本研究においては、個人的な観点での信頼性について着目する。

2.2 信頼情報を用いた推薦システム

信頼情報を利用した推薦システムに関して、さまざまな報告が行われている [9–12]。これらの研究では、ユーザ間のフォロー関係などのネットワークを基にして得られる、身の回りのユーザの信頼情報を考慮した推薦システムを提案している。本研究における提案モデルでは、身の回りのユーザに対する信頼情報を基に、ネットワーク上では遠く離れているユーザに対する信頼度を予測することができるという点で新規性がある。

3. 提案手法

本論文では、対象ユーザが既知のユーザに対する信頼度に基づいて、未知のユーザに対する信頼度を予測するモデルを提案する。提案モデルでは、対象ユーザは既知であるユーザに対する信頼度の評価を入力する。評価されたユーザの特徴と未知のユーザの特徴との間の類似度を求めることで未知のユーザに対する信頼度を推定する。なお、特徴量には、対象ユーザが信頼度を判断するための根拠となる情報を利用することが望ましい。本研究で対象とするオンラインレビューサイトにおける信頼判断の根拠は、レビューであると考え、今回はレビューを特徴量として利用する。

3.1 信頼推定モデル

3.1.1 対象ユーザのフィードバック

図1は、オンラインレビューサイトにおける信頼度評価の入力インターフェースの例を表す。対象ユーザはこのインターフェースを用いて、既知のユーザに対する評価を入力する。この図においては、レビュー文の右下に表示されている二つのアイコンが評価の入力ボタンであり、支持と書いてあるアイコンはポジティブな評価、不支持と書いてあるアイコンはネガティブな評価を入力するために利用する。また、レビューに対する評価がどちらも言えない場合は、どちらのアイコンもクリックしないことで、中立の表現とする。システムはユーザから入力された各レビューへの評価を、ユーザプロフィール情報として登録する。

3.1.2 既知のユーザに対する信頼度の推定

3.1.1で得られた評価情報に基づいて、対象ユーザが評価したユーザに対する信頼度を推定する。図2は、既知のユーザに対する信頼を推定するモデル図である。対象ユーザは既知のユーザが投稿した文書に対する評価を行う。対象ユーザ t の既知のユーザ u に対する信頼度 $\text{trust}(t, u)$ は、評価に基づき、以下の式で定義する。

$$\text{trust}(t, u) = \sum_{d \in D_u(\text{pos})} \text{typ}(u, d) - \sum_{d \in D_u(\text{neg})} \text{typ}(u, d) \quad (1)$$

この式では、対象ユーザからポジティブな評価を受けた文書集合の典型性の総和から、ネガティブな評価を受けた文書集合の典型性の総和を引くことにより、対象ユーザが既知であるユーザに対する信頼度を推定している。ここで、 $D_u(\text{pos})$ はユーザ u の投稿した文書集合のうち、対象ユーザからポジティブな評価を受けた文書集合を表し、 $D_u(\text{neg})$ はユーザ u の投稿した文書集合のうち、対象ユーザからネガティブな評価を受けた文書集合を表し、 d は1つの文書を表している。 $\text{typ}(u, d)$ は文書 d がどれだけユーザ u らしい文書であるかについての度合いを表し、本研究では、 $\text{typ}(u, d)$ をユーザ u における文書 d の典型性と定義する。典型性の詳細に関しては3.1.4で後述する。

3.1.3 未知のユーザに対する信頼度の推定

ここでは、3.1.2で得られた既知のユーザに対する信頼情報に基づいて、対象ユーザがまだ評価していないユーザに対する信頼度を推定する手法について述べる。本手法では、対象ユーザに評価されたユーザの特徴ベクトルと、まだ評価されていないユーザの特徴ベクトルとの間のコサイン類似度を求め、得られた値がまだ実際には評価されていないレビューに対する信頼度に影響を与える。対象ユーザ t の未知のレビュー u に対する信頼度 $\text{trust}(t, u)$ は、以下の式で定義する。

$$\begin{aligned} \text{trust}(t, u) = & \left(\sum_{d \in D(\text{pos})} \text{sim}(d, d_u) - \sum_{d \in D(\text{neg})} \text{sim}(d, d_u) \right) \\ & \times \text{typ}(u, d_u) \end{aligned} \quad (2)$$

この式では、対象ユーザからポジティブな評価を受けた文書集合と信頼を推定したい未知のユーザの文書との類似度の総和

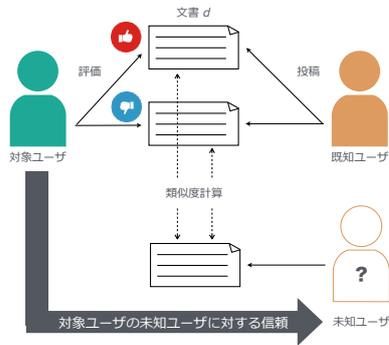


図 3 未知のユーザに対する信頼

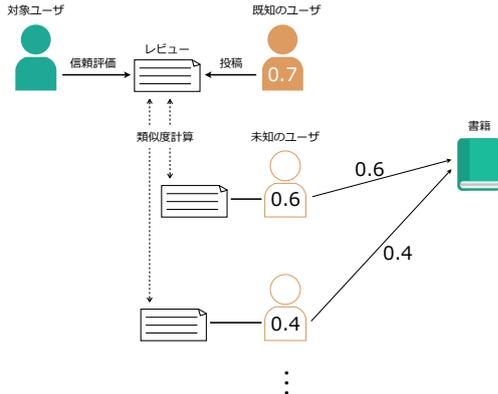


図 4 信頼を利用した推薦システム

からネガティブな評価を受けた文書集合と信頼を推定したい未知のユーザの文書との類似度の総和を引き、得られた値に文書の典型性を掛け合わせることで、対象ユーザが未知であるレビュワーに対する信頼度を推定している。ここで、 \mathbf{D}_{pos} は対象ユーザからポジティブな評価を受けた文書集合、 \mathbf{D}_{neg} は対象ユーザからネガティブな評価を受けた文書集合を表している。 $\text{sim}(d, d_u)$ は、文書 d と未知のユーザ u の文書 d_u との間の類似度を表し、Doc2Vec によりベクトル化された \mathbf{d} と \mathbf{d}_u の間のコサイン類似度を計算することで求める。

3.1.4 文書の典型性について

本研究では、評価された文書がどれだけそのユーザらしい発言であるかという典型性を信頼推定モデルの中に組み込むことで、計算される信頼度の確信性を向上させる。文書の典型性は、ユーザが今まで投稿した全ての文書をひとまとめにした文書のベクトルと、対象となる文書ベクトルとの間のコサイン類似度を計算することで求める。ユーザ u における文書 d の典型性 $\text{typ}(u, d)$ は以下の式で定義される。

$$\text{typ}(u, d) = \cos(\mathbf{d}, \mathbf{D}_u) \quad (3)$$

\mathbf{d} は典型性を求めたい文書のベクトルであり、 \mathbf{D}_u はユーザ u が投稿した文書全体のベクトルである。文書のベクトル化はどちらも Doc2Vec を利用して行う。

3.2 信頼度を利用した協調フィルタリング

提案モデルに基づき、オンラインレビューサイトにおけるレビュワーに対する信頼度を推定し、その値をレビュワーに対す

る重みとして、推薦書籍の計算および提示を行う。書籍 b_1 を選択した時の、ユーザ u に対する書籍 b_2 の予測評価値は、以下の式で求める。

$$\text{Pred}(u, b_1, b_2) = \sum_{r \in \mathbf{R}(b_1)} (\text{trust}(u, r) \times \text{read}(r, b_2)) \quad (4)$$

ここで、 $\mathbf{R}(b)$ は書籍 b に対してレビューを投稿したレビュワー集合を表し、 $\text{trust}(u, r)$ は対象ユーザ u にとってのレビュワー r の信頼度を表す。また、 $\text{read}(r, b)$ は、レビュワー r が書籍 b を読んだかどうかによって値が決まる。

$$\text{read}(r, b) = \begin{cases} 1 & (\text{読んだ}) \\ 0 & (\text{読んでいない}) \end{cases} \quad (5)$$

書籍 b_1 を読んだレビュワー集合のうち、書籍 b_2 を読んだレビュワーの信頼度の合計値がユーザ u に対する書籍 b_2 の予測評価値となり、これにより、ユーザの信頼度を考慮した協調フィルタリングによる推薦を行う。

3.3 Doc2Vec による文書のベクトル化

Doc2Vec は文書をベクトル化するための手法の一つである [13]。この手法は、語順を基にして単語の次の単語を予測できるような表現ベクトルを獲得する Word2Vec [14] を文章に拡張したものである。この手法により、Bag-of-Words などの従来手法では難しいとされる、文脈も含めたベクトル化を行うことが期待できる。Doc2vec によるベクトル化を行うためには、まず膨大な量の文書データをコーパスとして学習することが必要である。今回は、書籍レビューサイト「読書メーター」から取得したレビューを学習データとして用いた Doc2vec モデルを用いてレビュー文書のベクトル化を行う。

3.3.1 学習データの前処理

本節では、書籍を評価する際の評価内容をベクトルとして表現したい。そのため、書籍の具体的な内容に関する情報がベクトルの中に含まれるのは避けることが望ましい。そこで、書籍の登場人物名や作者名、など特定の書籍のみに出現するような単語を今回はストップワードとして扱い、学習の対象から除外するために、以下の前処理を行う。

全書籍のレビューを Mecab [15] を用いて形態素解析し、名詞・形容詞・形容動詞を抽出する。そして、それぞれの書籍ごとに、その書籍に対する全レビュー中の単語 w を含むレビューの割合を求める。形式的には、書籍 b における単語 w を含むレビューの出現率 $r(b, w)$ を以下の式で表す。

$$r(b, w) = \frac{|\{r \mid r \in \text{reviews}(b), w \in \text{words}(r)\}|}{|\text{reviews}(b)|} \quad (6)$$

ここで、 $\text{reviews}(b)$ は書籍 b を含むレビュー集合であり、 $\text{words}(r)$ はレビュー r を含む全単語からなる集合を表す。書籍集合 B における、それぞれの書籍に対する単語 w の平均出現率が低い場合は、その単語は特定の書籍固有の単語であると考えられる。書籍集合 B における単語の w の平均出現率 $r_{ave}(B, w)$ を以下の式で定義する。

$$r_{ave}(B, w) = \frac{\sum_{b \in B} r(b, w)}{|B|} \quad (7)$$

上記の式によって、求められた平均出現率 $r_{ave}(B, w)$ が設定した閾値より低い単語を学習の対象から除外する。平均出現率の低い単語が除外された、全レビューの全レビューを用いて Doc2Vec の学習を行う。ここで対象としたのは「読書メーター」に含まれる書籍レビューである。対象としたレビュー数は 9 万 5,719 人であり、これらのレビューの投稿した全てのレビュー、計 1233 万 1749 件を用いた。経験則に基づき、圧縮後の次元数は 300 とし、順番を考慮する単語幅は 8 と設定した。

4. 実験

プロトタイプを用いた被験者実験を行い、提案手法の有効性を評価した。

4.1 実験環境

オンライン書籍レビューサイト「読書メーター」[16]に登録しているレビューのうち計 9 万 5,719 人のレビュー、全 1233 万 1749 件を対象として、評価トピックを求めた。評価トピックの解析から除外する一般度の閾値は 0.35 とした。日本語解析には MeCab を用いて、辞書は mecab-ipadic-neologd [17] を使用した。ユーザに提示するレビューは、読書メーターに投稿されているレビューを用いた。

4.2 評価手法

対象とした被験者は 20 代の男女 8 人であった。被験者には、最近読んだ書籍の中から、自由に書籍を選んでもらい、その書籍のレビューに対して、評価のフィードバックを入力してもらった。評価については「このレビューを支持するか」という基準で「支持する」、「わからない」、「支持しない」の 3 段階の評価とし、「支持する」「支持しない」がそれぞれ 10 件以上入力されるまでを入力してもらった。プロトタイプシステムを通して得られた推薦結果上位 20 件のうち、未読書籍について「あらずじ」と「レビュー」を読んでもらい、「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の 5 段階でアンケート調査を行った。書籍の「1 巻」「2 巻」などは、まとめてシリーズとしての評価を行った。ベースラインは、信頼度を均一に設定した協調フィルタリングによる推薦とした。また、文書の典型性を考慮することの有効性を評価するために、典型性を考慮しない場合の提案手法による推薦も行った。

4.3 実験結果

「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の評価を、それぞれ 5 点、4 点、3 点、2 点、1 点として、手法ごとに、推薦書籍への評価の平均値を求めた。提案手法とベースラインのスコアを図 5 に示す。縦軸は推薦書籍への評価の平均値であり、横軸は左からベースライン、提案手法（典型性あり）、提案手法（典型性なし）である。また、評価値ごとの頻度分布を図 6 に示す。

4.4 考察

提案手法とベースラインとの間で、有意水準 5% において有

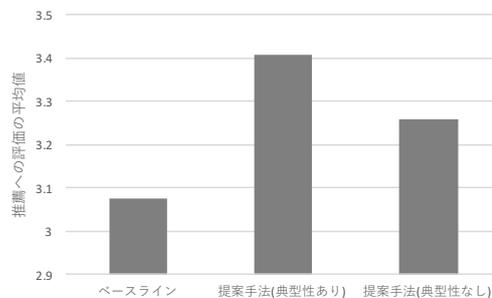


図 5 推薦書籍への平均評価値

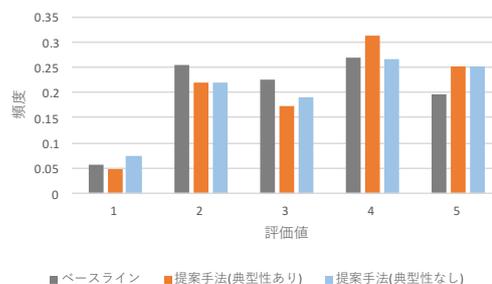


図 6 評価値の頻度分布

意な差が見られ、また、提案手法（典型性あり）はベースラインと比較して、評価値 1～3 の頻度が低く、反対に評価値 4～5 の頻度が高いことから、提案手法は書籍推薦の精度を上げるうえで有効であることが分かった。また、提案手法の中でも、典型性を考慮した手法としなかった手法を比較すると、典型性を考慮した手法の方が、考慮しなかった手法よりも、評価の平均値が高く、また評価値の頻度分布においても、評価値 4 の頻度が考慮しない手法より高く、評価値 1～3 の頻度は低かった。このことから、文書の典型性を考慮することが信頼度推定の精度を向上させる上で有効であることがわかった。

5. おわりに

本論文では、未知のユーザに対する信頼度を予測するモデルを提案し、提案モデルに基づいて書籍レビューサイトにおけるレビューに対する信頼度を予測し、得られた信頼情報を利用して書籍推薦を行なった。プロトタイプを利用した被験者実験の結果、提案モデルにより得られる信頼情報を利用した書籍推薦は、信頼情報を用いない従来型の推薦よりも、ユーザにとって満足度が高い推薦を実現可能であることが明らかになった。

今回提案した信頼度推定手法では、信頼度評価されたユーザとアイテムを介してノードが繋がっているユーザまでの信頼度推定を行なった。信頼度評価されたユーザとノードが繋がっていないユーザに対する信頼度は、アイテム間の類似度を利用して、推定できる範囲にいるユーザの信頼度を伝播させていくことで推定することが考えられる。今後は、ネットワーク上で遠く離れたユーザに対する信頼度推定や、SNS やシェアリングエコノミーなど、他の Web サービス上のユーザに対する信頼度推定に取り組む予定である。

謝 辞

本研究は（公財）電気通信普及財団の助成および JSPS 科研費 JP16K12534 の助成をうけたものです。

文 献

- [1] Airbnb, <https://www.airbnb.jp>
- [2] Uber, <https://www.uber.com>
- [3] TaskRabbit, <https://www.taskrabbit.com>
- [4] 伊木 淳, 亀井 清華, 藤田 聡, “レビューを対象とした信頼性判断支援システムの提案”, 情報処理学会論文誌, Vol.55, No.11, pp.2461-2475, 2014.
- [5] Mukherjee, A., Liu, B. and Glance, N.: Spotting Fake Reviewer Groups in Consumer Reviews, Proc. 21st International Conference on World Wide Web, pp.191-200 2012.
- [6] Xie, S., Wang, G., Lin, S. and Yu, P.S.: Review SpamDetection via Temporal Pattern Discovery, Proc. 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.823-831, 2012.
- [7] Wang, G., Xie, S., Liu, B. and Yu, S.: Review Graph based Online Store Review Spammer Detection, Proc. 11th IEEE International Conference on Data Mining, pp.1242-1247, 2011.
- [8] 小林 真雄, 伊藤 孝行, “電子商取引における関係性を考慮したユーザの信頼性評価”, 電子情報通信学会技術研究報告, Vol.109, pp.1-6, 2009.
- [9] H. Ma, I. King, and M. Lyu.: Learning to recommend with social trust ensemble. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp.203-210, 2009.
- [10] Simon Meyffret , Lionel Medini , Frederique Laforest, Trust-based local and social recommendation, Proceedings of the 4th ACM RecSys workshop on Recommender systems and the social web, 2012.
- [11] Bo Yang , Yu Lei , Dayou Liu , Jiming Liu, Social collaborative filtering by trust, Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, pp.2747-2753, 2013.
- [12] Allison J.B. Chaney , David M. Blei , Tina Eliassi-Rad, A Probabilistic Model for Using Social Networks in Personalized Item Recommendation, Proceedings of the 9th ACM Conference on Recommender Systems, pp.16-20, 2015.
- [13] Quoc V, Leand Tomas Mikolov. DistributedRepresentations of Sentences and Documents, In Proceedings of The 31st International Conference on Machine Learning, pp. 1188-1196, 2014.
- [14] Tomas Mikolov, Greg Corrado, Kai Chen, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space. In Proceedings of the International Conference on Learning Representations (ICLR 2013), pp.1-12, 2013.
- [15] <http://mecab.sourceforge.net/>
- [16] 読書メーター, <http://bookmeter.com>
- [17] <https://github.com/neologd/mecab-ipadic-neologd/blob/master/README.ja.md>