

レビューを対象とした信頼性判断支援システムの提案

伊木 惇 †

亀井 清華 ‡

藤田 聡 ‡

† 広島大学大学院 工学研究科

‡ 広島大学大学院 工学研究科

1 はじめに

近年、商品購入において、楽天市場などの口コミサイトのレビューは重要な役割を担っている。購買者は何の商品をどこで買うかといった選択をする際に、それらレビューを参考にすることが多いからである。そして、ポジティブなレビューであれば購買意欲を上げ、ネガティブな意見であれば購買意欲を下げる。

口コミサイトにおいて問題となるのは、サクラによるレビュースパムの投稿である。そのため、ユーザは投稿されたレビューの中から信頼できる情報であるか否かを判断する必要があるが、そのような信頼性を適切に判断することは難しい。理由として、**信頼性を判断するために役立つ情報が乏しく、加えて必要な情報を収集するコストが大きい**からである。上記問題を解決するために、本研究ではレビューの信頼性を判断するための支援システムの提案を行う。

本稿ではレビューの信頼性を表す指標を新たに定義する。そしてレビュー毎に各指標に関するスコアを提示することで、信頼性判断の支援を行う。

2 指標の提案

本研究で提案する4つの指標は以下の通りである。これらの指標はスパム検知に関する研究より得られた知見を基に提案する。基とした既存研究では、指標という形で提案されていない。そこで、本研究では信頼性の判断を支援するための指標を提案し、これらを用いた判断支援システムを構築する。

2.1 類似性

Jindai ら [1] は複製されたレビューには多くのスパムが含まれていることを示した。よって、どの程度他の文章と類似しているかを測る指標として**類似性**を提案する。

Jaccard 係数を用いてレビュー r_i の類似性スコアを以下のように求める。

$$S_score(r_i) = \max_{r_j} \left(\frac{|X_{r_i} \cap X_{r_j}|}{|X_{r_i} \cup X_{r_j}|} \mid j \neq i, j = 1, 2, \dots, n \right)$$

このとき、 n は r_i と同じジャンルに属するレビューの数である。また、 X_{r_i} と X_{r_j} は r_i と r_j を単語 bigram によって区切った文字を要素とする集合である。単語 bigram とは連続する2単語を1つの単位として区切る方法で

Credibility Judgment Support System for Review

†Makoto Igi ‡Sayaka Kamei ‡Satoshi Fujita

†Graduate School of SEngineering, Hiroshima University

‡Graduate School of SEngineering, Hiroshima University

ある。そして $|X_{r_i} \cap X_{r_j}|$ は r_i と r_j のどちらにも存在する要素数、 $|X_{r_i} \cup X_{r_j}|$ は r_i または r_j に存在する要素数を表す。

2.2 協調性

Mukherjee ら [2] は、同じ店の同じ商品に対して、高い(低い)評価を行うことによって、そのお店の評判を上げよう(下げよう)と協力し合っている投稿者グループが存在することを示した。どの程度他のユーザと協調的にレビューを行ったかを測る指標として**協調性**を提案する。

まず協力しあっているようなユーザグループを見つける。この時、頻出パターンマイニングにおける eclat アルゴリズム [3] を用いて、グループの大きさが3以上で、支持度(同じ商品にレビューした回数)が4以上となるような頻出ユーザグループ g_c を抽出する。そして、各 g_c の支持度(= $support(g_c)$) と大きさ(= $size(g_c)$) を掛け合わせたものを g_c の協調度($collaborate(u_{r_i}, g_c)_{u_{r_i} \in g_c}$) とする。 r_i の協調性スコアを以下のように求める。

$$C_score(r_i) = \ln(\max_{g_c \in G_{u_{r_i}}} (collaborate(u_{r_i}, g_c)))$$

このとき $G_{u_{r_i}}$ はユーザ u_{r_i} が属するユーザグループの集合である。

2.3 集中性

Xie ら [4] は、スパムは高い(低い)評価値を時間的に集中して投稿されることを示した。そのため高い(低い)評価値のレビューがどの程度時間的に集中しているかを測る指標として**集中性**を提案する。

どの程度レビューが集中しているかを求める方法として Kleinberg のバースト検知手法 [7] を用いる。これは時系列データに対してイベントの集中的な発生を検出する方法である。あるショップの評価値5のレビュー集合 $I = \{d_1, d_2, \dots, d_t\}$ を求め、1日刻みで区切る。そして、評価値5のレビューの割合が集中している日を求め、その日の集合を $I_{burst} \subseteq I$ とする。続いて集中している日 $d_{j,burst} \in I_{burst}$ の中で、再度 Kleinberg を用いて集中の原因となるレビューグループ $g_b \subseteq d_{j,burst}$ を求める。そして、そのグループの大きさをそのグループに属するレビューに対する集中性スコアとする。

$$T_score(r_i) = \begin{cases} 0 & (r_i \notin g_b) \\ \ln(size(g_b)) & (otherwise) \end{cases}$$

評価値1のレビュー集合に対しても同様の処理を行う。

2.4 情報性

Ott ら [5] は, レビューの特徴 (品詞や単語の意味) を用いてスパムかどうかの2値分類を行った. その中の知見として, informative かどうかでスパム検知できる可能性を示した. Rayson ら [6] によって informative な文章には名詞が多く使われていることがわかっている. そのため, どの程度 informative なレビューであるかを測る指標として**情報性**を下記のように定義する.

$$I_score(r_i) = \ln \left(1 + \sum_{j=1}^{|K_i|} \ln \left(\frac{n}{df(term_j)} \right) \right)$$

このときレビュー r_i に出現する名詞集合を K_i とし, $term_j \in K_i$ とする. $df(term_j)$ は r_i と同じジャンルのレビュー集合において $term_j$ を含んだレビューの数である.

3 実験

ここでは各指標に基づき得られたレビューを示す. 本実験では, 楽天市場の「みんなのレビュー・口コミ」データ*を用いる. 期間は2010年1月1日から31日で, 本(44668件), 家電(57469件), 家具(82064件)の3つのジャンルを用いる. また, レビューはユーザID・レビュー内容・評価値など全17項目が取得可能である.

3.1 実験結果

各指標に該当したレビュー件数を表1に示す. 対角線の値はそれぞれの指標に該当したレビューである. 類似性は $S_score = 1$ のレビュー (同じ文章が存在するレビュー) の件数, 協調性は $C_score > 0$ の件数, 集中性は評価値1または評価値5のレビューで $T_score > 0$ の件数, 情報性は $I_score = 0$ のレビュー (名詞を一切含まないレビュー) の件数である. また, 対角線以外の部分では, 両方の指標で該当したレビューの件数である. また, 3つの指標で該当したレビュー件数は6件 (類似性, 集中性, 情報性の場合のみ) であり, すべての指標で該当したレビュー件数は0件である.

また, 表2, 表3は表1の中から具体的に抽出したレビューの一例である. 表2は類似性スコアと集中性スコアが高く, 情報性スコアの低いレビューの例である. 明らかに複製して一度に投稿したレビューであることがわかる. そういった点において, スпамである可能性は高いと言える. また表3ではユーザaとユーザbのレビューに関する類似性スコアと協調性スコアが高い例である. 調査したところ, ユーザaは別アカウントのユーザcと文章が類似しており, ユーザbは別アカウントのユーザdと類似していた. また, ユーザaとbはすべて同じ

ショップの同じ商品にレビューを行っており, 評価値も高い. よって, お互いが協力し合って集中的にレビューを投稿することでショップの評判を上げている. そのため, このようなレビューもスパムである可能性が高い.

以上のように, 各指標を用いることで信頼性の低いと思われるレビューを確認することができる. よって, これらの指標を用いることでレビューが信頼できるかどうかの判断を支援することが可能だと考える.

【表1】抽出したレビュー件数

指標	類似性	協調性	集中性	情報性
類似性	9265	413	1141	131
協調性	-	1367	70	5
集中性	-	-	16372	65
情報性	-	-	-	813

【表2】抽出したレビュー例 (類似性, 集中性, 情報性)

店名	商品名	ユーザID	評価	レビュー本文	日付
店1	100円商品A	購入者さん	5	□□□□□□とっても可愛かったです□□□□□□	01-26 15:14:05
	100円商品B	購入者さん	5	□□□□□□とっても可愛かったです□□□□□□	01-26 15:14:38
	100円商品C	購入者さん	5	□□□□□□とっても可愛かったです□□□□□□	01-26 15:47:21

【表3】抽出したレビュー例 (類似性, 協調性)

店名	商品名	ユーザID	評価	レビュー本文	日付
店2	寝具A	aさん	4	肌当たる部分には優しい肌触りで, きっと赤ちゃんも大喜びです.	01-01 21:41:20
	寝具A	bさん	5	結構な冷え性の私なので, 早くも冷え対策の一環として購入しました.	01-16 23:01:39
	寝具B	cさん	5	肌当たる部分には優しい肌触りで, きっと赤ちゃんも大喜びです.	01-20 13:02:03
	寝具C	dさん	5	結構な冷え性の私なので, 早くも冷え対策の一環として購入しました.	01-16 23:27:57

4 まとめ

本研究では, レビューの信頼性判断支援のための指標の提案を行った. そして, 指標に基づき具体的に信頼性の低そうなレビューの例を示した.

参考文献

- [1] N.Jindai and B.Liu. "Opinion Spam and Analysis", *In Proceedings of WSDM 2008*, pages 219-230.
- [2] A.Mukherjee, et. al., "Spotting Fake Reviewer Groups in Consumer Reviews", *In Proceedings of WWW 2012*, pages 191-200.
- [3] M.J.Zaki, et. al., "New Algorithms for Fast Discovery of Association Rules." *In Proceedings of KDD 1997*, pages 283-286.
- [4] S.Xie, et. al., "Review Spam Detection via Temporal Pattern Discovery." *In Proceedings of KDD 2012*, pages 823-831.
- [5] M.Ott, et. al., "Finding Deceptive Opinion Spam By Any Stretch of the Imagination." *In Proceedings of HLT 2011*, pages 309-319.
- [6] P.Rayson, et. al., "Grammatical word class variation within the British National Corpus sampler." *Language and Computers 2002*, 36(1):pages 295-306.
- [7] J.Kleinberg. "Bursty and hierarchical structure in streams." *In Proceedings of KDD 2002*, pages 91-101.

*URL: <http://rit.rakuten.co.jp/rdt/>