

インターネットオークションにおける コメント文解析による評価指標の提案と応用

New Evaluation Criteria based on Comment Sentence Analysis in Internet Auction and its Application

植田嗣也*1
Tsuguya Ueta

伊藤孝行*1*2*3
Takayuki Ito

*1名古屋工業大学大学院産業戦略工学専攻

Master of Techno-Business Administration Nagoya Institute of Technology

*2マサチューセッツ工科大学スローン経営大学院
MIT Sloan School of Management

*3科学技術振興機構 (JST) さきがけ研究員
Researcher, JST PREST

Currently Internet Auction is one of the most important market for electronic commerce. However the existing evaluation systems are weak against dishonest behaviors, and they often fail to represent the correct evaluation values because they often use the single point evaluation system. In this paper, we present new evaluation criteria based on information from the review sentences in addition to evaluation values. As their application, we propose a method for finding dishonest users based on our new evaluation criteria. Our experimental results demonstrate that our method could find dishonest users at around 83 percent probability.

1. はじめに

インターネットの普及に伴い、インターネットを利用した電子商取引が盛んに行われるようになった。中でもインターネットオークションは、電子商取引の中でも大きな市場の一つである。取引を行う際、ユーザは、評価システムによって付けられた評価値や評価コメントなどを基に取引相手を判断することがあり、インターネットオークションにおける評価システムは非常に重要である。一方で、既存の評価システムには取引内容を数値として反映しきれていないという問題がある。単純に「良い」という評価が付けられたとしても、何がどの程度良いのかまでは数値的に反映されていない場合が多い。また、良い評価または悪い評価が反対の評価によって打ち消してしまうことも問題である。良い取引を行えば、与えられるポイントは1点であり、どんなに悪い取引を行ったとしても-1点しか与えられないのである。取引においては、評価の程度が異なるのが一般的であり、数値に評価の程度を反映させるべきである。

そこで本稿では Yahoo!オークション*1の取引データを解析し、コメント文解析に基づく新たな評価指標を提案する。具体的には、取引データから良い評価特有の単語と悪い評価特有の単語を求め、それらの単語の各評価コメントにおける出現回数と、コメントに伴う評価値から新たな評価指標を提案する。さらに応用として、評価指標に基づき信頼できないユーザの発見支援を行う手法を提案する。そして被験者実験を行い、アンケートを実施することで本手法の有効性を示す。

2. コメント文解析に基づく評価指標

2.1 概要

本章では Yahoo!オークションの取引データにおける評価コメント文を解析することにより、新たな評価指標 Point を提案する。Point の算出は、まず良い評価と悪い評価でのコメント

文を解析し、それぞれのコメント文における特有の単語を求め、その特有の単語を基に plus-words, minus-words, zerowords-plus および zerowords-minus を定義する。そして、これらの単語が評価コメント文にどのくらい頻出するかを調べ、頻出回数をカウントする。次に取引データにおける「非常に良い」「良い」「どちらでもない」「悪い」「非常に悪い」という評価に対して、それぞれ+2, +1, 0, -1, -2 の点数 (rate) を付ける。最後に rate と特有の単語の頻出回数から Point を算出する。次節から Point の算出方法、解析結果および取引データの取得方法について詳しく述べる。

2.2 コメント文の特徴を表す単語の選択

評価指標の提案において、良い評価と悪い評価を付けたコメント文の傾向を分析し、特徴を表す単語を求める。評価値が正と負のコメント文を、それぞれ mecab[1] を用いて形態素解析を行い、単語の頻出回数を調べる。そして頻出回数の多い単語から上位 30 語を plus-words および minus-words とする。ただし「ある」「この」などは上位にあったが特徴を表していないと考えられるので、このような単語は選択していない。また、plus-words および minus-words の中から「良い」「悪い」の評価を考えずとも、その単語だけでどちらの意味を表しているのかがわかる可能性の高い単語を 15 語選択し、それぞれ zerowords-plus および zerowords-minus とする。表 1 は特徴を表す単語の例を示す。「～ない」など、反対語による意味の反転は、コメント文にはほとんどなかった。そのため今回は反対語については考慮していない。

表 1: 特徴を表す単語の例

plus-words	ありがとう, 迅速, 丁寧, 商品, 信頼など
minus-words	残念, 遅い, 悪い, 連絡, 取引など
zerowords-plus	ありがとう, 迅速, 安心, 感謝など
zerowords-minus	残念, 遅い, 悪い, 返品など

2.3 評価指標の算出

まず、取引データのコメント文を評価値が正、負または 0 の 3 つに分類する。

連絡先: 植田嗣也, 名古屋工業大学大学院産業戦略工学専攻, 愛知県名古屋市昭和区御器所町, TEL:052-735-7968, FAX:052-735-7407, ueta@itolab.mta.nitech.ac.jp

*1 <http://auctions.yahoo.co.jp/jp/>

● 評価値が正のコメント文

コメント文における plus-words の頻出回数を調べ、これを count とする。そして評価値と count の積 (rate*count) を求め、これを新たな評価指標 Point とする。

● 評価値が負のコメント文

コメント文における minus-words の頻出回数を調べ、これを count とする。そして評価値と count の積 (rate*count) を求め、これを新たな評価指標 Point とする。

● 評価値が 0 のコメント文

zerowords-plus および zerowords-minus の頻出回数を調べ、それぞれを plus-count および minus-count とする。そして plus-count と minus-count の差 (plus-count - minus-count) を求め、これを新たな評価指標 Point とする。

2.4 取引データの解析

1. 評価の程度の数値化

Point の設定により評価値は、正の評価は+1 から+42、負の評価は-24 から-1、および 0 の評価は-5 から+10 までの範囲に分布される。特に評価値が正の評価では+2 から+20 まで、負の評価では-1 から-8 まで、および 0 の評価では-1 から+6 までの範囲に多くの Point が分布される結果となった。良い評価の中でもより良い評価、悪い評価の中でもより悪い評価を付けることができ、既存の評価システムよりも評価の程度の違いを数値的に反映することができたと考えられる。以下にコメント文とそれに対する Point の例を示す。

- Point : +16 (評価 : 非常に良い)
このたびは最後まで迅速で丁寧なご対応をいただきましてありがとうございました。とても信頼ができる出品者様です。また機会がありましたら宜しくお願いします。
- Point : +8 (評価 : 非常に良い)
無事着いて安心いたしました。とても信頼できる落札者です。スムーズにお取引できました。またどうぞご利用下さい。
- Point : -8 (評価 : 非常に悪い)
何度もメールしましたが一度も連絡いただけませんでした。購入する意思がないのなら入札しないでいただきたいです。非常に迷惑です。
- Point : -16 (評価 : 非常に悪い)
落札後なかなか連絡ももらえず、こちらから連絡をしてやっと連絡ををもらったかと思えば、商品がない。最悪ですね。ない商品は出品しないでください。非常に迷惑です。
- Point : +4 (評価 : どちらでもない)
本日商品届きました。迅速・丁寧な対応ありがとうございました。また機会がありましたら宜しくお願いします。
- Point : -2 (評価 : どちらでもない)
連絡は遅かったし商品発送も遅かった。

2. 評価指標のばらつき

Point の設定により、設定前にはほとんどなかった点数のばらつきを大きくさせることができた。表 2 は評価値が正、負および 0 の 3 つの場合における Point の設定前と設定後の変動係数である。変動係数とは標準偏差を算術平均で割った値であり、単位のない数となり、相対的なばらつきを表すことができる。この値が大きいほど、ばらつきが大きいといえる。

Point を設定することにより、変動係数は $rate > 0$ の評価では約 10 倍、 $rate < 0$ の評価では約 4 倍増加していることがわかる。また $rate = 0$ の評価ではすべて 0 ポイントであった評価に対して、ばらつきをつけることができた。これにより、良い評価、悪い評価およびどちらでもないという評価の中でそれぞればらつきが大きくなり、差を広げることができた。

表 2: 変動係数の比較

	Point 設定前の 変動係数	Point 設定後の 変動係数
$rate > 0$ の評価	0.04700641	0.51920228
$rate < 0$ の評価	-0.20204290	-0.84774540
$rate = 0$ の評価	ゼロ除算	0.87235429

2.5 取引データの取得

本研究ではクローラーを作成し、Yahoo!オークションの評価ページからデータの取得を行った。クローラーは Ruby で実装し、データベースに MySQL を使用した。取得方法はまず、探索基点となる Yahoo!オークション ID を入力する。ID を受け取ったクローラーは Yahoo!オークションの評価ページから HTML を取得する。次に、取得した HTML を解析し、オークション ID や今までの総合評価、コメント文などを取得して MySQL に格納する。そしてクローラーは評価ページでの評価者の Yahoo!オークション ID を次の探索候補とし、取引データを取得していく。表 3 に取得した取引データの各パラメータおよび、図 1 にクローラーの概念図を示す。ここで rep の値は、過去の評価の合計値を表している。「非常に良い」または「良い」の評価をされれば+1、「どちらでもない」の評価なら 0、「悪い」または「非常に悪い」の評価なら-1 が反映され、それらの値の合計値が rep となる。ただし、同じ人と何度も取引をして、そのたびに評価を受けた場合、総合評価に反映されるのは、最新の評価のみとなっている。また rate の設定では、評価が「非常に良い」= +2、「良い」= +1、「どちらでもない」= 0、「悪い」= -1 および「非常に悪い」= -2 の 5 段階で評価値を設定している。取得した取引データ数は評価回数が 3,728,407 回、owner name が 9,765 人、および traded name が 1,871,028 人となっている。

表 3: 取引データ

Field	Type	説明
owner_name	varchar	対象とするページのユーザ ID
traded_name	varchar	取引相手の ID
rep	int	取引相手の総合評価
rate	int	取引相手のその取引での評価値
comment	text	取引相手のその取引でのコメント

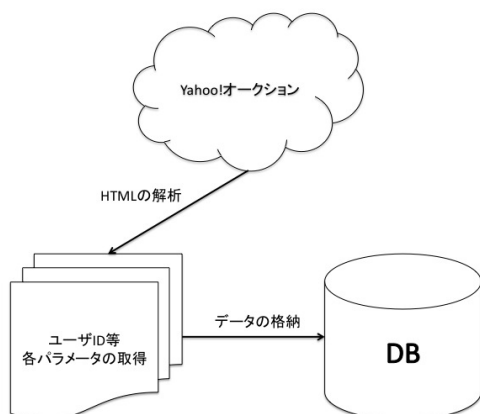


図 1: クローラーの概念図

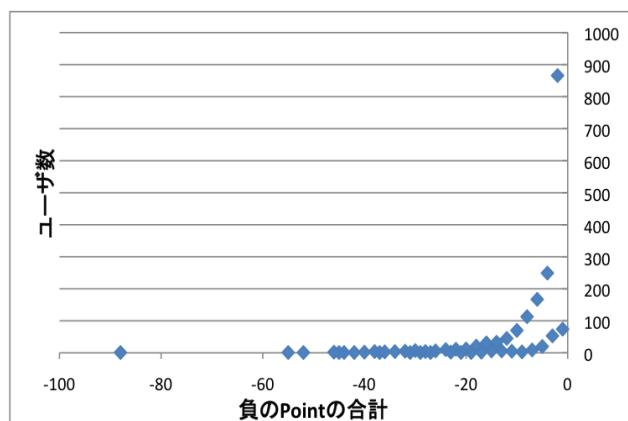


図 2: ユーザごとの負の Point の散布図

3. 信頼できないユーザの発見支援

提案した評価指標の応用として、信頼できないユーザの発見支援を行う手法を提案する。以下の 2 つの条件に当てはまるユーザに注目する。

1. 高評価ユーザによる悪いコメント文の多いユーザ

高評価ユーザによる悪いコメント文と、そうでないユーザによる悪いコメント文では信頼度に差があると考え、高評価ユーザによって付けられた負の Point を分析する。取得した取引データから合計評価値が平均以上であるユーザを高評価ユーザとし、各ユーザごとに高評価ユーザから付けられた負の Point を求める。この値が低いユーザは、取引経験が多く、良い評価を得ているユーザから悪いと判断された取引を多く行ったユーザである。図 2 に高評価ユーザからつけられた負の Point をユーザごとに示す。

2. 悪い内容のコメント文が多いユーザ

Point と評価値の差を利用し、評価値だけではわからない、悪い内容のコメント文を多く書かれているユーザの抽出を行う。ここで負の評価において、Point と評価値の差 (合計 Point - 合計評価値) を gap と定義する。gap が低いユーザは、ある程度悪い評価を付けられ、かつ悪い内容のコメント文が多いユーザであると考えられる。ユーザごとの gap の散布図を図 3 に示す。

上記の 2 つの手法を評価するためにアンケートを行い、詳しい結果を次章で述べる。

4. 評価

本手法を評価するために 20 代の男女、合計 20 人を対象としてアンケートを行った。図 4 に結果を示す。2 つの手法において、1 の手法については負の Point の合計値が低い 11 ユーザ、2 の手法については gap の合計値が低い 10 ユーザをそれぞれ抽出し、その中から主観により選択した 18 ユーザに関してアンケートを行った (図 4 の u1 ~ u18)。内訳は 1 の手法から 10 ユーザ (図 4 の u1 ~ u10)、2 の手法から 8 ユーザ (図 4 の u11 ~ u18) を選択した。回答者には各ユーザについて信頼

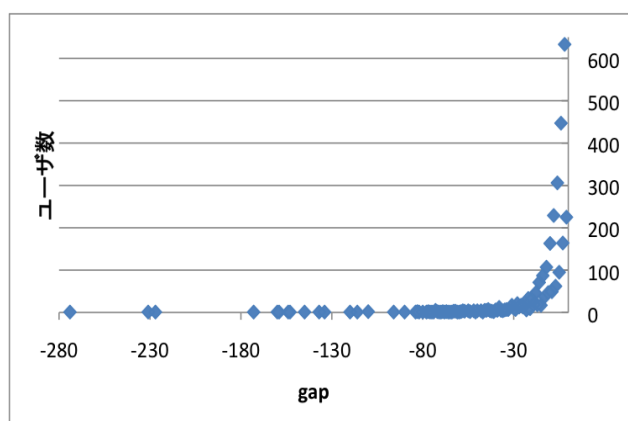


図 3: gap の散布図

できるかどうか回答させ、またユーザを判断する上で最も大事だと思う判断要素も回答させた。

図 4 において丸で囲まれているユーザは、出品者、落札者のどちらか 1 つの立場でも「信頼できない」という回答の方が「信頼できる」という回答よりも明らかに多かったユーザである。(1) の手法での抽出ユーザに関しては 10 ユーザ中、7 ユーザ、(2) の手法での抽出ユーザに関しては 8 ユーザ中、5 ユーザとなった。また、丸で囲まれていないユーザは「信頼できる」「信頼できない」の回答数が同じくらいであったユーザ、または「信頼できる」の回答数が明らかに多かったユーザである。提案手法は負の Point に注目した手法であり、比較的悪い評価が多いユーザを抽出している。しかし、抽出したユーザの中には、悪い評価の何倍も良い評価を付けられているユーザが存在したためこのような結果になったと考えられる。全体として、信頼できるとはいえないユーザは全体の約 83%(15/18)であった。以上より、本手法により信頼できないユーザを発見する支援ができたといえる。

またユーザを判断する上で最も大事だと思う判断要素は、9 割以上の回答者が総合評価やコメント文と回答した。提案した評価指標は、評価値とコメント文に基づいた Point であり、多くの人が判断要素として重要と考えている要素を考慮した評価指標を提案することができた。

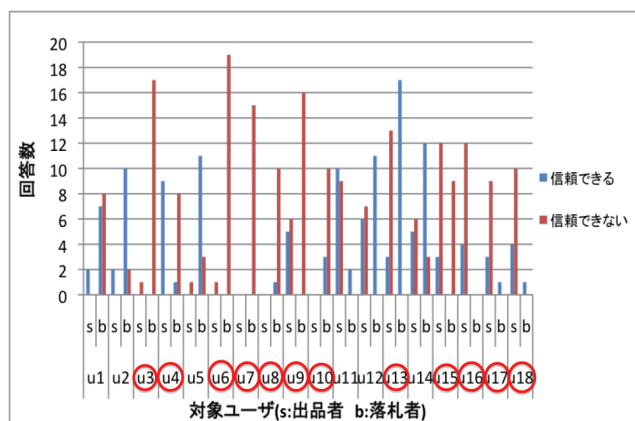


図 4: アンケート結果

5. 関連研究

5.1 インターネットオークションにおける評価指標

小林 [2][3] らは、インターネットオークションサイトをネットワーク構造として捉え、ネットワーク構造を評価する Auction Network Trust を提案している。そして Auction Network-Trust をオークションネットワークに適用することで、個々のユーザの出品者、落札者としてのスコアを算出している。また新井 [4] らは、Yahoo!オークションの情報から、ユーザが行おうとしている取引の安全性を、点数と信頼度が感覚的に合うよう数値化し、取引のリスクを明確化することによってオークションに参加する際の意思決定を支援するシステムを提案している。

5.2 評判分析

立石 [5] らは、ユーザが入力した商品名とあらかじめ辞書として用意した評価表現を近接演算する方法を用いて、インターネットの Web ページから意見を抽出している。また、抽出した意見の適性値を構文的な特徴を利用して判定している。藤村 [6] らは、Web 掲示板から評価を抽出し、分類している。Web 掲示板では、書き込みそのものに肯定や否定のタグが付与されているものがある。この情報を用いて、肯定の評判に現れる単語と否定の評判に現れる単語を学習し、評価表現の分類を行っている。分類は単語のスコア計算をすることで、肯定と否定に分類している。肯定的 (否定的) な評判には、肯定的 (否定的) な概念を持った語が多く含まれているはずであるという仮定のもと、肯定的な評判と否定的な評判の差をとることで、スコアリングをしている。大野 [7] らは、インターネットオークションの評価コメントを、記述した落札者に注目して要約する方法を提案している。落札者ごとにその人の記述した出品者評価コメントを調べ、多くの出品者に対して記述しているような儀礼的なコメント文は省き、特定の出品者に対して記述してある思い入れがあると思われるようなコメント文を抽出している。

本研究では、藤村 [6] らと同様の仮定のもと、「非常に良い」「良い」評価のコメント文と「非常に悪い」「悪い」評価のコメント文に分け、それぞれから評価における特有の単語を求めた。また、新井 [4] らのコメント文を基にした評価関数においても、特有の単語を求めている。しかし彼らの研究ではサンプル数が十分ではないと考えられる。また「度」「もの」「何」などのような特徴を表しているとは考えられない単語まで考慮してしまっているが、本研究ではこのような特徴を表していないと考えられる単語は省略した plus-words, minus-words など

を設定している。

6. まとめ

本稿ではインターネットオークションにおける評価のコメント文を基にした評価指標 Point を提案した。Point の提案により既存の評価システムよりも評価の程度の違いを数値的に反映することができたと考えられる。良い評価の中でもより良い評価に高いポイントを、悪い評価の中でもより悪い評価に低いポイントを、どちらでもないという評価にはどちらかというといと良いと判断できる評価に正のポイント、悪いと判断できる評価に負のポイントをそれぞれ付けることができた。また既存の評価システムのポイントにはほとんどなかったばらつきを大きくさせることができた。Point の提案によってユーザは取引相手として考えているユーザの数値を見ることで、ある程度そのユーザの評価を知ることができるようになった。また Point の応用として、信頼できないユーザの発見支援を行う手法を提案した。アンケート評価の結果から、約 83% の割合で信頼できるとはいえないユーザの発見をすることができた。この結果から、信頼できないユーザの発見を支援することができたといえる。

今後の課題としては、まず、評価指標において単語ごとに重み付けをすることが考えられる。例えば「遅い」という単語よりも「詐欺」という単語の方がより悪い意味を持つと考えられ、重みを付けることで精度の向上を目指す。また本稿では取引データにおいて、コメント文に対する返答までは考慮していないため、今後は返答まで考慮した評価指標の提案を考えていきたい。さらに評価指標を -10 から +10 までのように上限と下限を設定することで、ユーザにとってわかりやすい評価指標になると考えられる。

参考文献

- [1] MeCab: Yet Another Part-of-Speech and Morphological Analyzer
<http://mecab.sourceforge.net/>
- [2] 小林真雄, 安藤哲志, 伊藤孝行”電子商取引ネットワークにおけるユーザ評価指標の提案”, 共同エージェントワークショップシンポジウム (JAWS2008), p.681-688, 2008
- [3] 小林真雄, 伊藤孝行”電子商取引におけるユーザの関係性に関するクラスタリング手法の提案”, 電子情報通信学会, 信学技報, vol. 109, no. 211, AI2009-9, pp. 1-6, 2009
- [4] 新井純一, 高橋貞夫, 大倉典子”インターネットオークションにおける意思決定支援システム”, 電子情報通信学会, 電子情報通信学会技術研究報告. SSS, 安全性 pp.9-12, 2007
- [5] 立石健二, 石黒義英, 福島俊一”インターネットからの評判情報検索”, 人工知能学会誌, Vol.19, No3, pp.317-323(2004)
- [6] 藤村滋, 豊田正史, 喜連川優”電子掲示板からの評価表現および評判情報の抽出”, 人工知能学会全国大会 (第 18 回)(2004), <http://www-kasm.nii.ac.jp/jsai2004-schedule/pdf/000192.pdf>
- [7] 大野華子, 楠村幸貴, 土方嘉徳, 西田正吾”社会的関係を用いたネットオークション評価コメントの自動要約”, 電子情報通信学会論文誌 (D), Vol.J-88-D-I, no.3, pp.668-683, March 2005