

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新しい評価指標の提案とその応用

植田 嗣也^{†1} 伊藤 孝行^{†1,†2}

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。現在広く用いられている評価システムには、不正行為に悪用されたり、評価内容を数値的に反映しきれていないという問題がある。そこで本論文では、インターネットオークションサイトから取引情報を収集し、取引後の評価値とコメント文の解析により新しい評価指標を提案する。また応用として、評価指標に基づいて不正行為ユーザの発見支援を行う手法を提案する。評価指標の提案により、評価の程度の違いを数値的に反映することができた。さらにアンケートによる評価から、提案手法では約83%の確率で信頼できるとはいえユーザの抽出をすることができた。

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

TSUGUYA UETA^{†1} and TALAYUKI ITO^{†1,†2}

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} 名古屋工業大学: Nagoya Institute of Technology

^{†2} MIT スローン経営大学院: MIT Sloan School of Management

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と考えられる。しかし、評価システムが重要になってくる一方で、既存の評価システムには取引内容を数値的に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。また応用として、評価指標に基づき不正行為ユーザの発見支援を行う手法を提案する。そしてアンケートによる評価を行うことで提案する手法の有効性を示す。

2. 関連研究

2.1 インターネットオークションにおける詐欺の発見

NetProbe²⁾ は、米カーネギーメロン大学のコンピュータサイエンス研究者グループが開発したインターネットオークション上の詐欺を発見するシステムで、オークションの取引ネットワークから特殊な構造を発見するというものである。この研究チームは詐欺者同士の評価値のつり上げを行った取引履歴が、2部コアグラフを構成することに着目し、そこから詐欺グループを同定しようとしている。

平手³⁾らは、インターネットオークションの評価システムにおいて、評価値つり上げによる急激な評価値の増加を検出することで詐欺グループの抽出を行う研究をしている。この研究では、評価値の時系列変化を調べ、短期間に大きく増加し、その後急激に増加が収まるようなIDを異常として検出している。また、この手法で抽出されたIDの取引履歴から不正評価コミュニティを抽出している。

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

2.2 インターネットオークションにおける評価指標

小林⁴⁾⁵⁾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⁶⁾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意思決定を支援するシステムを提案している。

2.3 評判分析

立石⁷⁾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⁸⁾らは、Web 掲示板から評判を抽出し、肯定と否定に分類している。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、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。大野⁹⁾らは、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。

3. コメント文解析に基づく評価指標

3.1 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。

図 1 にクローラ概念図、表 1 に取得した取引データの各パラメータ、表 2 に取得した取引データ数を示す。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」「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」「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。

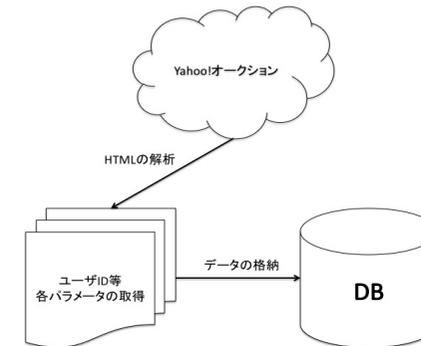


図 1 クローラーの概念図

表 1 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

表 2 取得したデータ数

評価回数	3,728,407 回
owner_name	9,765 人
traded_name	1,871,028 人

3.2 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、まず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 mecab¹⁾ 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」などは上位にあって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 3 は特徴を表す単語の例を示す。

表 3 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

3.3 評価指標の算出

図 2 に提案する評価指標 Point の算出方法を示す。まず、取引データのコメント文を評価値が正、負、0 の 3 つに分類する。

- 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。

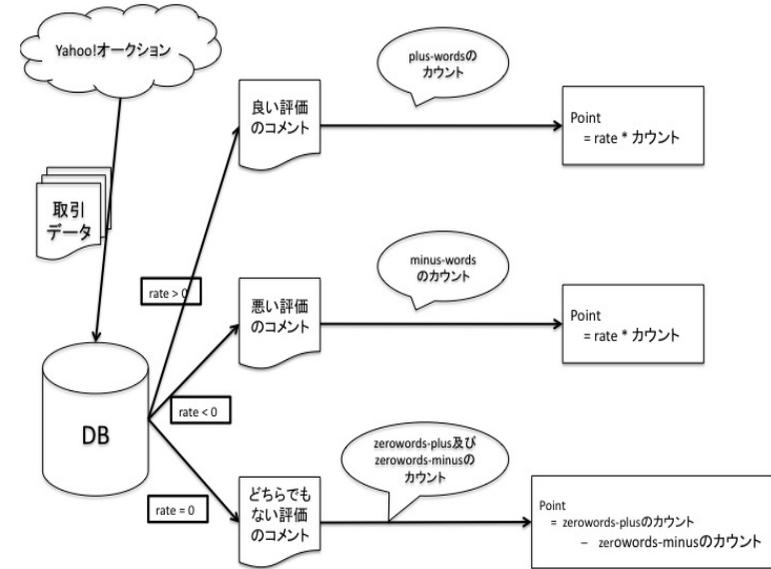


図 2 評価指標 Point の算出方法

3.4 評価指標の解析

(1) 評価の程度の数値化

Point の設定により、評価値が正の評価は+1 から+42, 負の評価は-24 から-1, 0 の評価は-5 から+10 までの範囲に分布された。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 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ので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

表4は評価値が正、負、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。

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、ばらつきをとけることができた。これにより、良い評価、悪い評価、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4 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

4. 不正行為ユーザの発見支援

提案した評価指標の応用として、不正行為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3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4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。

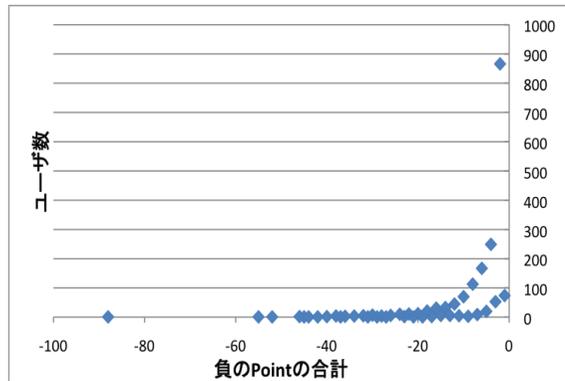


図3 ユーザごとの負の Point の散布図

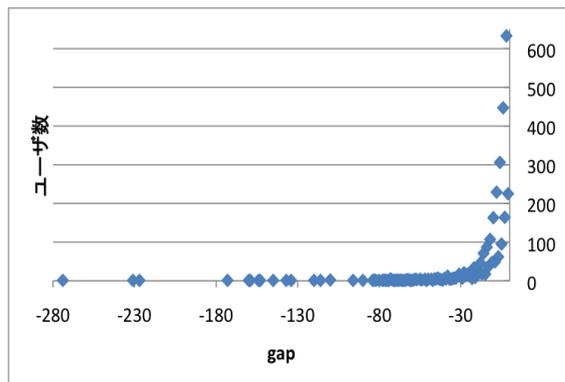


図4 gap の散布図

5. 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5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5のu1~u18)。内訳は(1)の手法から10ユーザ(図5のu1~u10)、(2)の手法から8ユーザ(図5のu11~u18)

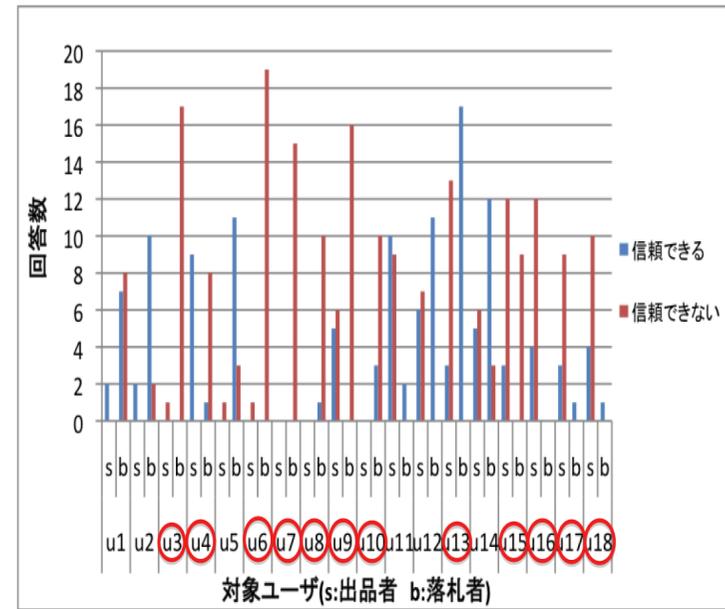


図5 アンケート結果

を選択した。回答者には各ユーザについて「信頼できる」「信頼できない」「五分五分」「判断できない」のように4段階評価で回答させ、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5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。

6. まとめと今後の課題

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、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ように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。これにより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のではないかと考えられる。また Point の応用として、不正行為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ので、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、ユーザは取引相手の候補として考えている相手の評価の数値を見ることでより判断しやすくなることが考えられる。

参 考 文 献

- 1) MeCab : Yet Another Part-of-Speech and Morphological Analyzer
<http://mecab.sourceforge.net/>
- 2) Shashank Pandit, Duen Horng Chau, Samuel Wang, Christos Faloutsos, Net-Probe: A Fast and Scalable System for Fraud Detection in Online Auction Networks, World Wide Web, 2007
- 3) 平手勇宇, 相吉澤明, 翁松齡, 井奥雄一, 木戸冬子, 山名早人, "インターネットオークションにおける不正行為者の発見支援," 日本データベース学会 Letters, Vol.5, No.2,

pp. 77-80, 2006.

- 4) 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
- 5) 小林真雄, 伊藤孝行"電子商取引におけるユーザの関係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
- 6) 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
- 7) 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)
- 8) 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>
- 9) 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