

# 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<sup>1,a)</sup> 亀井 清華<sup>2,b)</sup> 藤田 聡<sup>2,c)</sup>

受付日 2014年3月13日, 採録日 2014年9月12日

**概要:** ec サイトにおける商品のレビューは、商品購入の意思決定に大きく関わり、価値ある情報として注目されている。一方で、ステルスマーケティングを目的とした、レビュースパムと呼ばれる信頼性の低いレビューの投稿が問題となっている。既存研究では、レビューの文章などから、それらスパムを検知する取り組みが行われてきた。しかしながら依然として、すべてのスパムの検知は難しい。さらに、レビューを読むユーザ自身が判断するにも、信頼性を判断するための情報は十分でない。また、ユーザは、ウェブ上の情報に対して、ある程度信じやすいという報告もされている。そのため、ユーザが信頼性を意識し、判断するための機構が必要である。よって、本稿では ec サイトにおけるレビューを対象とした信頼性を判断するための支援システムを提案する。具体的には、レビューの信頼性を表す指標として、類似性、協調性、集中性、情報性という4つの信頼性指標を定義し、各指標ごとのスコアを求める。そして、レビューごとにそのスコアを可視化して提示する。それにより、ユーザ自身に信頼性を意識してレビューを読むように促すとともに、信頼性判断がしやすくなるよう支援を行うことが可能となる。本研究では、これらの指標を用いた判断支援を行うシステムを構築し、評価を行った。その結果、提案システムにより、ユーザの信頼性に対する意識を促すとともに、有効な判断支援が行えることが確認できた。

キーワード：レビュー、スパム、信頼性、電子商取引サイト

## Credibility Judgment Support System for Review

MAKOTO IGI<sup>1,a)</sup> SAYAKA KAMEI<sup>2,b)</sup> SATOSHI FUJITA<sup>2,c)</sup>

Received: March 13, 2014, Accepted: September 12, 2014

**Abstract:** Reviews of products in e-commerce sites such as Rakuten have attracted as valuable information. On the other hand, in such the sites, unreliable reviews called review spam have become a big issue. In existing works, they proposed various methods to detect the spam. However, spam detections play a cat-and-mouse game with new type of spam, and any spam detections are not enough for the issue. Therefore, for users, mechanisms to support judgments of the credibility of each review are necessary. Thus, we proposed a support system to judge the credibility for reviews in e-commerce sites. Specifically, we define four credibility indicators to represent how much each review is spammy. Then, our support system calculates scores for each indicator and provides the scores for users. In this paper, we built a prototype system and evaluated the system by questionnaires. As a result, by using our system, it was confirmed that it is possible to enhance awareness of credibility for users.

**Keywords:** review, spam, credibility, e-commerce

<sup>1</sup> 中国電力株式会社  
The Chugoku Electric Power Co., Inc., Tottori 680-0812, Japan

<sup>2</sup> 広島大学大学院工学研究院  
Department of Information Engineering, Hiroshima University, Higashihiroshima, Hiroshima 739-8527, Japan

a) makoto.igi.en@gmail.com

b) s-kamei@se.hiroshima-u.ac.jp

c) fujita@se.hiroshima-u.ac.jp

## 1. はじめに

近年、Amazon<sup>\*1</sup>や楽天市場<sup>\*2</sup>などの電子商取引サイト(以下 ec サイト)におけるレビューは重要な役割を担っている。ウェブ上の商品購入者の60%は、評価値やレビュー

\*1 URL: <http://www.amazon.co.jp/>.

\*2 URL: <http://www.rakuten.co.jp/>.

を含んだサイトから商品を購入する傾向にあり、購入者の70%以上は、購入前にレビューを読んで参考にしていることが報告されている [1]。またユーザは、レビューがポジティブであれば、その商品に対する購買意欲を高め、ネガティブであれば、購買意欲を下げる。このように、投稿されたレビューの内容がユーザや店側の金銭上の利益へとつながる。そのため、レビューは価値ある情報としてマーケティング分野 [2] など様々な分野から注目されており、レビューの文章に対する言語処理やデータマイニング技術を用いた研究 [3], [4], [5], [6] もさかんに行われている。

しかしながら、これらの ec サイトではその利益を目的としたサクラによるステルスマーケティングがしばしば問題となっている。サクラとは商品や店の評判を上げる（または下げる）ために、レビュースパムという不当な偽のレビューの投稿を行う悪意を持った投稿者のことである。レビュースパムは、そのレビューを読んだユーザを騙し、誤解を促す。場合によってはユーザや店が被害を被る可能性もある。さらに、ec サイト上ではレビューの売買<sup>\*3</sup>が行われており、スパムへの対策は差し迫った問題とされている [7]。そのため、各レビューがスパムであるか否かを評価し、判断するための機構が必要だと考えられる。本稿では、この「スパムであるか否か」あるいは「投稿者に悪意があるか否か」がそのレビューの信頼性を表すと考え、あるレビューがスパムである可能性（スパムらしさ）が高いことを信頼性が低いとする。また、仮にレビュースパムを読んだユーザが商品を購入して満足したとしても、そのレビュー投稿者がサクラと疑わしい行動をとっていけばスパムらしいとする。

これまでに、レビューの信頼性に関して様々な研究 [7], [8], [9], [10], [11] が行われてきた。各レビューがサクラによって投稿されたレビューであるか否かを正確に判断することは不可能である。そのため、これらの研究ではサクラやスパムが持つであろう特徴を示し、そのような特徴をもつ投稿者やレビューの検知手法の提案をしている。しかしながら、これらの研究は、それぞれがスパムの一部についての一面的な特徴のみをとらえたものであり、さらなる改善の必要性が述べられている [8]。さらに、スパムとその取り組みとの間にはたちごっこ側面があり、これらの手法を組み合わせたとしても十分だとはいえない。そのため、最終的には人の主観に頼る必要があると考えられる。これらの既存研究においても、各検知手法の精度の評価の際には、各レビューが信頼できるか否かを複数の人手によって判断してもらい、その判断結果との比較を行っている。

一方で、以下の理由により、それぞれのレビューが信頼できるか否かを人手によって判断することは難しい。

- (1) レビューのソース、証拠情報といった信頼性を判断するために役立つ情報が乏しい [9], [12]。
- (2) 信頼性判断のために必要な情報を個人が十分に収集するにはコストが大きい [12]。

よって、本稿では上記問題を解決するために、レビューの信頼性に関する指標をユーザに示すことで信頼性判断の支援を行うシステムを提案する。本稿では、レビューの信頼性を表す指標として、類似性、協調性、集中性、情報性という4つの信頼性指標を定義する。これらの指標はレビュースパムに関する既存研究 [7], [8], [9], [10] において議論されたスパムのよく知られた特徴に基づいて提案するものであり、各レビューの「スパムらしさ」を示すものである。既存研究がその特徴にマッチするものをスパムであるとして検知するのに対し、本研究ではこれらの特徴にマッチする度合いをユーザに示すことにより、最終判断はユーザにゆだねるものの、それにより判断を容易にすることを目指している。また、ウェブ上の情報に対してユーザはある程度信じやすいことが報告 [13] されているが、これらの指標を提示することで、ユーザ自身が信頼性を意識してレビューを読むように促すことが可能となる。本研究では最後に、ユーザに対するアンケートにより、システムの評価を行う。スパムであるかないかの判断が容易になっているかどうかは、既存研究 [7] で行われたように、複数の人手による評価において意見が1つにまとまるかどうかで評価する。その結果、提案システムによって、ユーザに対して信頼性に対する意識を促せること、また、有効な判断支援が行えることが確認できた。

本稿の構成は以下のとおりである。2章で関連研究について述べる。3章では本研究で新たに定義した指標を示す。各指標の基となったスパムが持つ特徴についてもここで紹介する。4章で、提案システムの概要を示す。5章では、システムの評価結果と考察を述べる。最後に6章でまとめと今後の課題を述べる。

## 2. 関連研究

本章ではまず、レビューの信頼性に関するレビュースパムの研究を紹介する。その後、ウェブページに対する信頼性判断支援システムに関する研究を紹介する。

### 2.1 レビュースパム検知

Jindai らによる文献 [8] は、レビュースパムに関する最初の研究である。そこで彼らはスパムには3つのタイプが存在することを示している。

- Type1 (untruthful opinions)：商品の評判を上げる（または下げる）ことを目的として、ポジティブな（またはネガティブな）レビューを不当に投稿することで読み手やシステムに誤解を与えてしまうレビュー
- Type2 (review on brands only)：特定の商品に対する

<sup>\*3</sup> <http://www.nytimes.com/2011/08/20/technology/finding-fake-reviews-online.html>

レビューを行うのではなく、その製造者や販売店に対してのみ言及しているレビュー

- Type3 (non-review) : 広告, または意見を含まないレビュー (質問, 解答, ランダムなテキスト)

Type2 と Type3 はレビューの文章から容易に判断できる. そのため Type1 のスパムを検知することが課題となっている. しかしながら, その課題を解決するにあたって大きな問題がある. それはどのレビューが本当にスパムであるかという答えとなる情報が存在しないことである. そこで Jindai らは, 複製されたレビューに着目した. 複製, もしくは複製に近いレビューを収集し, レビューの文章を読むことでそれらがスパムとして妥当かどうか, 人手による検証を行った. 結果, 複製, もしくは複製に近いレビューには Type2 と Type3 が多く含まれていることが分かった. また, 複製, もしくは複製に近いレビューの中で Type2, Type3 ではないものの中には, 以下の3つものが多く含まれていた.

- 異なるユーザ ID で同じ商品に投稿された同じ (あるいは非常に似た) 本文のレビュー
- 同じユーザ ID で異なる商品に投稿された同じ (あるいは非常に似た) 本文のレビュー
- 異なるユーザ ID で異なる商品に投稿されているが同じ (あるいは非常に似た) 本文のレビュー

これらは Type1 である可能性が高いといえる. そのため, 複製されたレビューをスパムとして考えるのは妥当だとした. そして, 複製されたレビュースパムを検知する手法の提案を行っている. しかしながら, すべてのスパムの検知には至っていないことを述べている.

Mukherjee ら [7] は, 協調的にレビューを投稿することで商品あるいは店の評判を上げる (または下げる) 投稿者グループがいると考えた. これは, 複数の投稿者 (サクラ) が特定の商品または店に対して, 高い (または低い) 評価で多くのレビューの投稿を行うことで, その商品または店の評判を上げる (または下げる) ことが可能であり, そのような投稿を行うサクラグループが存在するという事である. そこで, Mukherjee らは, データマイニングの分野における頻出アイテムセット抽出の方法を用いて, サクラグループの候補となる投稿者グループ (ユーザ ID の集合) の作成を行った. そして, 作成したいくつかの投稿者グループに対して, 8 人の専門家 (ec サイトの従業員) に, サクラグループであるかどうかの判断をしてもらった. その結果, サクラグループであるかどうかの意見が多くの専門家同士で一致した. つまり, 投稿者グループが与えられた場合であれば, それがサクラグループであるかどうかの判断が容易となることが分かった. これは, 同じグループに属する投稿者間の投稿履歴や, 投稿されたレビューの文章間の比較ができ, それらの情報がスパムであるかどうかの良い判断材料となったからである.

前述のように, 一般的に単一のレビューから, サクラ, もしくはスパムであるかどうかの判断はユーザにとって難しい [8]. しかしながら, Mukherjee らのように, あらかじめ投稿者の投稿履歴やレビューの文章の比較を可能にすることで, サクラグループであるかどうかの判断が容易になることが示された. つまり, システム側であらかじめスパムらしさの情報を抽出し, それを判断材料として示すことでユーザのスパムの判断が容易になることが期待される.

## 2.2 ウェブページの信頼性判断支援システム

Yamamoto ら [12] は, ウェブページを対象とした信頼性判断支援システムを提案している. 彼らのシステムでは, ウェブページに関する信頼性判断指標として5つの指標を定義し, 各指標に関するスコアをウェブページの横に表示することで判断支援を行っている. Yamamoto らは, ウェブ検索結果の信頼性をユーザに意識させ, ユーザが自身の信頼性判断基準に従ってウェブ検索結果の信頼性を判断しながら, 最終的に信頼性の高いウェブページを取得するための支援を目的としている. Yamamoto らはウェブページを対象として Accuracy (コンテンツの参照重要度), Authority (コンテンツの社会的受容度), Objectivity (コンテンツの客観性), Coverage (専門的なトピックの網羅量), Currency (コンテンツの鮮度と更新頻度) の5つの指標を提案し, それぞれのスコア化を行った.

しかしながら, これらをレビューの信頼性指標としてそのまま用いることはできない. たとえば, Accuracy に関するスコアを求める方法として, Yamamoto らは Google によって提供されている PageRank を用いた. しかし, ec サイトにおいてはレビューどうしのリンク関係は存在せず, PageRank を求めることができない. また, Objectivity に関するスコアを求める方法としては, はてなブックマークのブックマーク数を用いている. ec サイトでは, 「参考になった」ボタンが提供されている場合もあるが, その場合でもほとんどのレビューで「参考になった」ボタンが押されておらず, また「参考になった」ボタン自体多くのバイアスがかかっていることが報告されている [5]. よって, これをレビューの信頼性指標に用いるにはあまり有用でないと考えられる. このように, レビューとウェブページとは扱うことのできる情報が異なっており, そのままの指標を用いることができない. よって, 本研究では Yamamoto らとは異なる, レビューのための新たな指標を提案する.

ec サイトでのレビューには, 金銭的やりとりが発生するため信頼性判断の支援の必要性が一般のウェブページにおけるそれよりも高い. そのため, 本研究では, 複数の信頼性指標をユーザに提示することによって信頼性判断の支援を行うことを目的とする. このように, レビューに対して, ユーザによる判断を支援する研究は, 我々の知るところではこれまでになされていない.

### 3. 信頼性指標の提案

本章では、本システムで用いる信頼性指標に関する定義を行う。これらの指標はスパム検知に関する既存研究より得られた知見を基に提案する。また、各指標のスコアは0以上5以下の値をとるスコアへと正規化する。このとき、値が5に近いほどスパムであることが疑われるものとする。

#### 3.1 類似性

Jindai ら [8] は複製またはそれに近い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。そこで、どの程度、他のレビューの文章と類似しているかを測る指標として類似性スコアを提案する。

ここでは、文献 [14] の考えを利用して文章の類似度を測る。まず、レビュー  $r_i$  の文章を単語 bigram により区切る。これは接続する2単語を1つの単位要素として区切る方法である。これによって区切られた単位要素の集合をレビュー  $r_i$  を表す要素集合  $X_{r_i}$  とする。

次に、Jaccard 係数を用いてレビュー  $r_i$  と  $r_j$  の類似度を以下のように求める。

$$\text{sim}(r_i, r_j) = \frac{|X_{r_i} \cap X_{r_j}|}{|X_{r_i} \cup X_{r_j}|}$$

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

そしてレビュー  $r_i$  の類似性スコアを以下のように求める。

$$S\_score(r_i) = \max_{r_j} (\text{sim}(r_i, r_j) \mid j \neq i, j = 1, 2, \dots, n)$$

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。そして、下記の方法によりスコアを正規化する。

$$S\_score_{norm}(r_i) = 5 \cdot S\_score(r_i)$$

#### 3.2 協調性

Mukherjee ら [7] は、商品の評判を上げる（または下げる）ことを目的とした、サクラグループの存在を明らかにした。これは同じグループのメンバが同じ商品に対して投稿を行い、協力して評判を変えるものである。そこで、各レビューがサクラグループによって投稿されたものである可能性を測る指標として協調性スコアを提案する。

まずサクラグループを見つけるために、頻出アイテムセット抽出の方法 [7] を用いる。まず  $t_{p_i}$  を、ある商品  $p_i$  にレビューを投稿したユーザ ID の集合とし、トランザクションと呼ぶ。また、ある投稿者グループが出現したトランザクションの数をそのグループの支持度数と呼ぶ。そして、支持度数が4以上でユーザ ID の数が3以上となる頻出投稿者グループ  $g_c$  を求める。今回、計算には統計解析

ツール R<sup>\*4</sup>における eclat アルゴリズム [15] を用いた。

次に、各  $g_c$  の支持度数 (=  $\text{support}(g_c)$ ) とユーザ ID 数 (=  $\text{size}(g_c)$ ) を用いて  $g_c$  の協調度を以下のように計算する。

$$\text{collaborate}(g_c) = \text{support}(g_c) \cdot \text{size}(g_c)$$

そして、レビュー  $r_i$  の協調性スコアを以下のように求める。

$$C\_score(r_i) = \begin{cases} \ln(\max_{g_c \in G_{u_{r_i}}} (\text{collaborate}(g_c))) & |G_{u_{r_i}}| \neq \emptyset \\ 0 & |G_{u_{r_i}}| = \emptyset \end{cases}$$

このとき  $u_{r_i}$  は  $r_i$  を投稿した投稿者であり、 $G_{u_{r_i}}$  は  $u_{r_i}$  が属する頻出投稿者グループの集合である。

さらに、下記の方法によりスコアを正規化する。

$$C\_score_{norm}(r_i) = \frac{5 \cdot C\_score(r_i)}{\max(C\_score(r_j) \mid j = 1, 2, \dots, N)}$$

このとき  $N$  はすべてのレビューの数である。ただし、投稿履歴が公開されていない投稿者に関してはスコアを求めることはできない。

#### 3.3 集中性

Xie ら [9] や Zhang ら [16] は、レビュースパムが時間的に集中して投稿されることを示した。そこで、各店のレビューに対して、高い（または低い）評価値のレビューがどの程度集中して投稿されているかを測る指標として集中性スコアを提案する。評価値とは、各レビュー投稿時にレビューの文章に添えて投稿される5段階の値である。

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 [17] 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。たとえば、特定の「単語」を含んだウェブページの投稿が急激に増えることがある。そのような現象をバーストと呼び、Kleinberg の手法はこれを検出する用途に用いられている。ここでは「単語」を「評価値」に置き換えることで、特定の評価値のレビューの増加度合いを検知し、そのときのその評価値を持つレビューの集中性スコアとする。

Kleinberg のバースト検知手法 [17] には、単位時間ごとに発生したイベントを数えあげた離散的な時系列データに対する離散型バースト検知手法と、刻々と発生する連続的な時系列イベントに対する連続型バースト検知手法とがある。離散型手法は、「単位時間ごとのイベントの数でだんより割合が増えているとき」を検知する。よって、たとえば1日ごとのレビュー数を計測して用いるのには有効であるが、より細かい時間的な変化を観測するために短い

\*4 URL: <http://www.r-project.org/>.

単位時間での計測を行おうとすると、前後の単位時間との違いが上手く計測できない場合があり、単位時間の設定が難しい。一方の連続型手法は、「イベント間の時間間隔がふだんより短くなっているとき」を検知する。よって、これだけを用いても、まったく投稿のない日と投稿のあった日の差は検出できるが、その差に引きずられて1日のうちの細かい変化などを観測するのは難しい。そこで我々はこれらの2つを組み合わせ、ある評価値の投稿数の割合が急激に増えた日を離散型手法で検知した後、その日の中でのその評価値の投稿数の時間的な変化を連続型手法で計測することとした。以下に、評価値5のレビューの集中度スコアを求める方法を説明する。

ある店の  $m$  日目のレビュー集合を  $B_m$  とし、時刻の早い順から  $B_1, B_2, \dots, B_m$  と離散時間で送られてくることを考える。このような  $m$  日分のレビュー集合に対して、Kleinberg の離散型バースト検知手法を用いて、ふだんよりも評価値5のレビューの割合が増えている日を求める。そのような日を仮に  $t$  日目とし、 $t$  日目の評価値5のレビュー集合を  $B_{t_5}$  とする。

次に、 $B_{t_5}$  の要素を投稿された時間順に並べた投稿時間列  $\mathbf{r} = \{r_1, r_2, \dots, r_{u+1}\}$  を考える。そして、 $r_j$  と  $r_{j+1}$  の投稿時間間隔を  $x_j$  としたとき、 $\mathbf{r}$  の投稿時間間隔列  $\mathbf{x} = \{x_1, x_2, \dots, x_u\}$  を求める。そして、投稿時間間隔列  $\mathbf{x}$  に対して Kleinberg の連続型バースト検知手法を用いることで、投稿時間間隔が連続して短いレビュー集合  $g_b \subseteq B_{t_5}$  を求める。各レビュー  $r_i \in g_b$  の集中度スコアは、 $g_b$  のレビューの数  $size(g_b)$  を用いて以下のように求める。

$$T\_score(r_i) = \ln(size(g_b))$$

ただし、どのレビュー集合  $g_b$  にも属さないレビューの集中度スコアは0とする。さらに、下記の方法でスコアを正規化する。

$$T\_score_{norm}(r_i) = \frac{5 \cdot T\_score(r_i)}{\max(T\_score(r_j) \mid j = 1, 2, \dots, N)}$$

本システムでは同様にして、評価値1のレビューの場合についてもスコアを求める。

### 3.4 情報性

Johnson ら [18] は、文章が informative であるほど、スパムではない可能性が高いことを示した。informative であるとは、有益な情報を多く含んでいることをいう。また、Rayson ら [19] によって、informative な文章は、名詞が多く使われていることが分かっている。つまり、レビュー本文中に特徴的な名詞の割合が少ないほどスパムらしいといえる。Ott らはこれらのことを使って、スパム検知手法を提案している [10]。そこで、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スコアを以下のよ

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  $n$  はレビュー  $r_i$  と同じジャンルに属するレビューの数である。また、 $r_i$  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とする。よって、同じジャンルの中でもあまり他のレビューでは使われていないような特徴的な名詞を多分に含んだレビューであればスコアが高くなる。そして、下記の方法でスコアを正規化する。

$$I\_score_{norm}(r_i) = 5 \cdot \left( 1 - \frac{I\_score(r_i)}{\max(I\_score(r_j) \mid j = 1, 2, \dots, n)} \right)$$

ここで、 $I\_score(r_i)$  自体はスパムでない可能性を表しており、 $I\_score_{norm}(r_i)$  は他の指標によるスコアと同様にスパムらしさを表していることに注意されたい。

### 3.5 予備実験

上記で提案した4つの指標に基づきレビューを抽出する。ここでは各指標ごとにスコアが5となる場合のレビューの件数と4以上となる場合のレビューの件数を求める。

#### 3.5.1 データセット

本実験では、楽天市場の「みんなのレビュー・口コミ」データ\*5を用いる。期間は2010年1月1日から31日で、本(44,668件)、家電(57,469件)、家具(82,064件)の3つのジャンルを用いる。また、各レビューはユーザID・レビュー内容・評価値など全17項目が取得可能である。

#### 3.5.2 レビューの抽出

各指標に該当したレビュー件数を表1に示す。“()”の外の値は各指標のスコアが5となる場合のレビューの件数であり、“()”の中の値はスコアが4以上となる場合のレ

表1 抽出したレビュー件数

Table 1 The number of extracted reviews by scores.

指標	件数
類似性	9,265 (10,191)
協調性	67 (448)
集中度	59 (213)
情報性	813 (1,930)
類似性+協調性	3 (273)
類似性+集中度	7 (123)
類似性+情報性	131 (276)
協調性+集中度	0 (0)
協調性+情報性	0 (3)
集中度+情報性	0 (5)

\*5 URL: <http://rit.rakuten.co.jp/rdr/>.

ビューの件数である。スコアが5の場合、3つの指標で該当したレビュー件数は0件であった。スコアが4以上の場合では、類似性・集中性・情報性の組合せの場合に限り、3件該当した。スコアが5の場合も4以上の場合も、4つすべての指標に同時に該当した件数は0であった。

また、注目すべき点として、

- 類似性スコアの高いレビューが多い点、
  - 協調性と集中性の両方が高いレビューが存在しない点、
- があげられる。類似性スコアの高いレビューが多い理由は、一度に似たような商品を複数購入した場合に、同じ文章をコピーして投稿する投稿者の存在が考えられる。たとえば、コミックでは、ある作品の1巻と2巻を同時に購入することがある。この場合、同じ作品のため、同じ内容のレビューを投稿することが考えられる。実際に調査したところ、類似性スコアが5であったレビューの内訳は、本(5,438件)、家電(1,497件)、家具(2,330件)であり、本のジャンルのレビューに多く見られた。

次に、協調性と集中性の両方が高いレビューが存在しない理由について考える。サクラグループは同じ期間にレビューを投稿すると考えられるため、協調性スコアと集中性スコアの高いレビューが投稿されるはずである。しかし、集中性スコアは、「ある店において、評価値5(または1)のレビューが“ふだんよりも”多く投稿されている期間の評価値5(または1)のレビューは怪しい」とするものである。そのため、ふだんから評価値5(または1)のレ

ビューが多く投稿されている店のレビューには、高い集中性スコアはつかない。今回の場合、協調性スコアの高いレビューが投稿された店に対するレビューのうち、評価値5のものが約7割を占めていた。そのため、協調性と集中性の両方が高いレビューは観測されなかった。

表2、表3は表1の中から具体的に抽出したレビュー(家具ジャンル)の一例である。表2は類似性スコアの高いレビューの例である。これらのレビュー本文中の「□□□□□」は、元のレビューにおいてもこの表記で、装飾的に用いられているものである。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分かる。そういう点において、スパムである可能性は高い。また表3ではユーザaとユーザbのレビューに関する協調性スコアが高い例である。このことはレビュー本文と評価値からは分からないが、ユーザaとbについて調査したところ、表4のようなレビューが確認された。aとbはすべて同じ店の同じ商品にレビューを投稿しており、評価値も高い。また、投稿した日付も一致している。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店の評判を上げている可能性が高い。そのため、このようなレビューもスパムである可能性は高い。

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た。よって、これらの指標を用いてレビューの信頼性の判断を支援するシステムの提案を行う。

表2 抽出したレビュー例(類似性スコア=5)

Table 2 Examples of the extracted reviews (The similarity score is 5).

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(協調性スコア=5)

Table 3 Examples of the extracted review (The cooperativeness score is 5).

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	肌に当たる部分には優しい肌触りで、きつと赤ちゃんも大喜びです。	01-01 21:41:20
	寝具B	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

表4 調査したレビュー

Table 4 Other reviews by users a and b of Table 3.

店名	商品名	ユーザID	評価	レビュー本文	日付
店2	寝具B	aさん	5	肌触りもサラリとして気持ちいいです。このお値段でこの品質なら文句なしです。	01-16 23:30:52
	寝具A	bさん	4	お返しに購入しましたが、大変喜ばれました。とても暖かいみたいです。	01-01 21:28:57
	寝具C	aさん	5	価格も安くお値段的に十分です。またよろしくお願いします。	01-16 23:32:14
	寝具C	bさん	4	これだけの値段でこの素材なので満足してます。	01-16 23:05:14
	寝具D	aさん	4	腰痛がひどいため、マットでためしたいとおもいます。良くなるといいなあ・・・	01-19 14:17:22
	寝具D	bさん	5	明日届く予定なのですが、レビューみてるので早く試したいです。	01-19 13:55:39
	寝具E	aさん	5	寒い夜には、羽毛布団が一番。今回で2つ購入。家族で愛用中。	01-19 14:19:43
	寝具E	bさん	4	羽毛布団初購入、今日から暖かい毎日がまっています。	01-19 13:58:50

### 4. 提案システム

レビューの信頼性を判断するには、レビューの内容だけでなく、他のレビューやユーザの投稿履歴など、様々な情報の中から判断する必要がある。しかしながら、ユーザが読むレビューすべてに信頼性の判断を行おうとすると、膨大な時間的コストがかかってしまう。また、判断材料となる情報にたどりつくことさえ難しい。そのため、判断を容易にするための機能やその判断材料となる情報を提示することが必要である。よって、本システムは、下記の2つの機能を提供する。

(1) レビューに対するスコアの可視化

(2) スコアの根拠となる情報の提示

これらの機能により、ユーザの信頼性判断を容易にし、支援することが可能となる。

(1) では、3章で述べた指標に関するスコアを可視化する。具体的には、Yamamoto ら [12] と同様にレーダーチャートを用いる。図1がレビューに関する信頼性指標を可視化した例である。このように可視化することで、ユーザに対して、レビューの信頼性を意識させるとともに、様々な指標からレビューの信頼性に関して直感的に判断させることができる。

(2) では、各指標に関する情報を提示する。具体的には、以下の情報へのアクセスを可能とするリンクを、各レビューのレーダーチャートと共に提示する(図2参照)。

類似性リンク：類似性は、どの程度他のレビューの文章と

類似したレビューかを測る指標である。そのため、類似する他のレビューにアクセスするためのリンクを提示する。それにより、ユーザは、どのようなレビューと類似しているかを確認することができる。このとき、リンク先には、最も類似したレビューのみ提示する(図3の③参照)。最も類似したレビューが複数あれば、それらすべてを提示し、1つしかなければその1つを提示する。

協調性リンク：協調性は、そのレビューがサクラグループによって投稿された可能性を測る指標である。そのため、同じ頻出投稿者グループ  $g_c$  (3.2節参照) に属する他のユーザ情報(図4の④参照)にアクセスするためのリンクを提示する。また、その際にどの程度同じ商品にレビューを投稿しているかを表す投稿者間の類似度を示す。このときの類似度は、0以上1以下の値をとり、1が最も類似していることを表す。さらに、同じ商品へのレビュー内容を確認するためのリンク(図4の⑤参照)も提示する。図4では投稿者のユーザIDが user23796 のレビューの協調性指標の根拠として④のグループが提示され、その中の各投稿者に関する情報として、その投稿者が user23796 と同じ商品に対して投稿したレビュー⑤を示している。

ユーザの履歴情報がなく、協調性スコアが測定できない場合には、「測定不能」と表示する(図2の②参照)。

集中性リンク：集中性は、どの程度時間的に集中して投稿されたレビューかを測る指標である。そのため、時間的に近い他のレビューにアクセスするためのリンクを提示する。それにより、ユーザは、どのようなレビューと投稿時間が近いかを確認することができる。このとき、リンク先には、連続して投稿されたレビュー集合  $g_b$  (3.3節参照) のレビューを提示する(提示方法は、

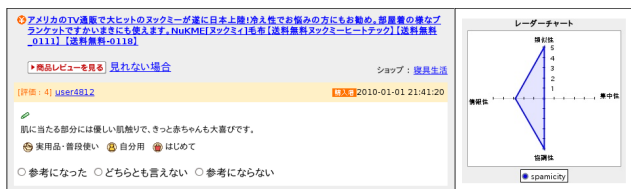


図1 スコアの可視化

Fig. 1 The visualization of scores.

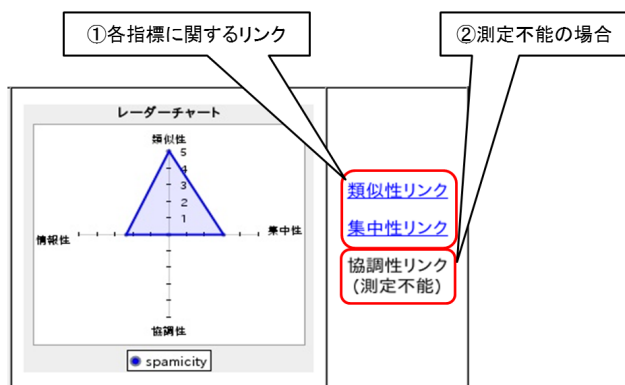


図2 各指標に関する情報へのリンク

Fig. 2 The hyperlinks to the information about the reasons of scores.

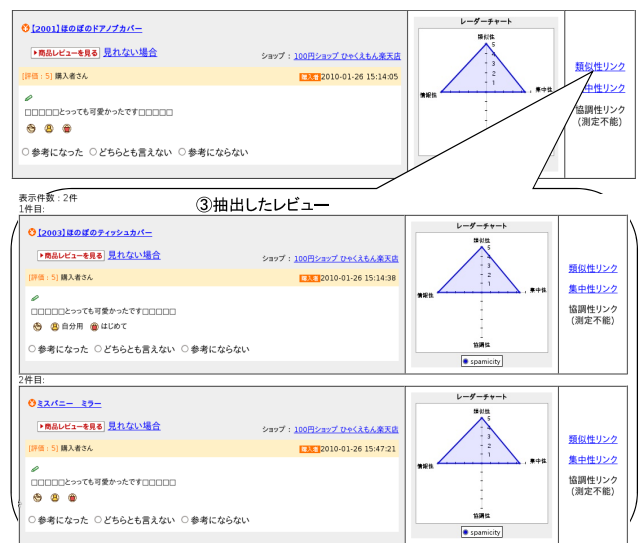


図3 リンク先(類似性)

Fig. 3 The hyperlink direction about similarity score.

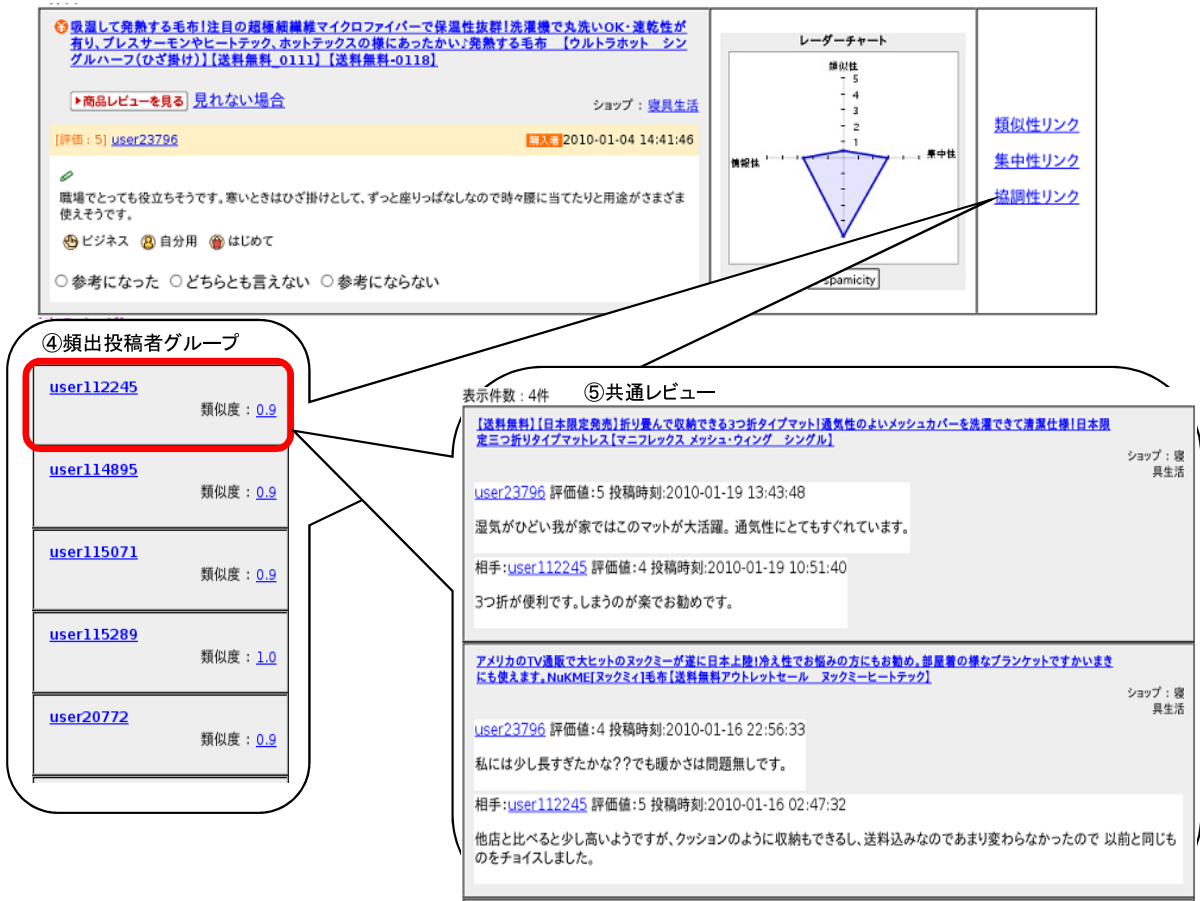


図 4 リンク先 (協調性)

Fig. 4 The hyperlink direction about cooperativeness score.



図 5 システム画面

Fig. 5 The system interface.

類似性のリンク先 (図 3) と同様の形式)。

以上のように、各指標に関する情報を提示する。これらのリンクにより、ユーザが判断材料とするための情報へのアクセスが容易になり、ユーザは、直接判断することが可能となる。ただし、情報性は、同じジャンルの他のレビュー全体であまり使われていない単語をもとに計算されているが、他の個々のレビューとの直接の関係を示す指標ではないという点で他の指標とは異なるため、情報性に関するリンクは提示していない。

#### 4.1 プロトタイプシステム

上記の2つの機能を用いることで、ユーザのレビューに対する信頼性判断支援を行うためのシステムを構築した。システム画面を図5に示す。サーバー環境は、Java 1.6.0.27, Apache-Tomcat 7.0.47, SQLite 3.7.8を用いた。本プロトタイプシステムでは、3.5.1項で述べたデータセットを用いて、各指標のスコアについてあらかじめ計算しておくものとした。本システムの実用化を考えた場合、各スコアについて、定期的に計算しなおすことにより、同様の効果が得られるものとする。

### 5. 実験

提案システムの有効性の検証を行う。方法としては、提案システムを利用しない(指標やリンクを見せない)場合と、利用する(指標やリンクを見せる)場合のレビューの信頼性に対するアンケートを行う。そして、アンケートの結果から、提案したシステムの有効性の評価を行う。

#### 5.1 事前実験

まず事前実験として、ecサイトを利用する際の意識調査を行った。被験者は、大学生・大学院生21人とする。質問項目は以下のとおりである。



表 5 事前アンケート結果の内訳 1 [人数 (割合)]

Table 5 The result of the preliminary questionnaire 1.

ふだんからレビューがスパムであるかどうかは意識しますか？	
はい	9 (43%)
いいえ	12 (57%)

- (1) ふだんからレビューがスパムであるかどうか意識しますか？
- (a) はい
  - (b) いいえ
- (2) ある商品を購入する場合、意識的に見る情報は何ですか？ (複数回答可)
- (a) 商品情報
  - (b) 商品全体の評価値\*6
  - (c) 商品のレビュー数
  - (d) 商品に関するレビュー
  - (e) 商品を提供している店の他の商品情報
  - (f) 商品を提供している他の店の商品のレビュー
  - (g) 同じジャンルの異なる商品情報
  - (h) 同じジャンルの異なる商品に関するレビュー
  - (i) その他
- (3) あるレビューを読む場合、意識的に見る情報は何ですか？ (複数回答可)
- (a) レビューの文章
  - (b) レビューの評価値
  - (c) レビューの投稿時間
  - (d) レビューの投稿者履歴
  - (e) レビューが購入者によって投稿されているか
  - (f) その他

質問項目 (2) はある商品を購入する際に、その商品のどのような情報を見るかの意識調査であり、質問項目 (3) は、ある商品のあるレビューを見ている際に、そのレビューのどのような情報を見るかの意識調査である。

表 5, 表 6 に、上記の質問項目に対する回答結果を示す。表 5 より、大半の被験者がふだんからスパムを意識していないことが分かる。また、表 6 では、スパムをふだんから意識する人としない人に分けてそれぞれにおける割合を示している。この結果より、スパムを意識する被験者であっても、他の商品のレビューや、レビューの投稿時間、レビューの投稿者履歴などの情報はあまり見ないことが確認される。しかし、スパムを意識するうえでは、これらは重要な情報である。そのため、信頼性判断の意識を促すこと、また判断を支援する必要性は高いといえる。

## 5.2 実験方法

実験は、21 人の大学生、大学院生に対して行った。まず、提案システムを利用しない実験を行った。本実験では、

\*6 評価値 1 のレビューが何件、評価値 2 のレビューが何件といった商品に対する評価値の分布情報。

表 6 事前アンケート結果の内訳 2 [人数 (割合)]

Table 6 The result of the preliminary questionnaire 2.

ある商品を購入する場合、意識的に見る情報は何ですか？		
	スパムを意識する人	スパムを意識しない人
商品情報	9 (100%)	12 (100%)
商品全体の評価値	7 (78%)	6 (50%)
商品のレビュー数	4 (44%)	6 (50%)
商品に関するレビュー	8 (89%)	10 (83%)
商品を提供している店の他の商品情報	0 (0%)	0 (0%)
商品を提供している店の他の商品のレビュー	0 (0%)	0 (0%)
同じジャンルの異なる商品情報	4 (44%)	7 (58%)
同じジャンルの異なる商品に関するレビュー	4 (44%)	3 (25%)
その他	0 (0%)	0 (0%)
レビューを読む場合に、意識的に見る情報は何ですか？		
	スパムを意識する人	スパムを意識しない人
レビューの文章	9 (100%)	11 (92%)
レビューの評価値	7 (78%)	11 (92%)
レビューの投稿時間	1 (11%)	0 (0%)
レビューが購入者によって投稿されているか	4 (44%)	3 (25%)
レビューの投稿者履歴	1 (11%)	0 (0%)
その他	0 (0%)	0 (0%)

被験者に対してレビューを 1 件ずつ提示し、被験者に、そのレビューが信頼できるかどうかのアンケートに答えてもらった。このとき、レビュー内容に加えて、各種情報 (投稿者履歴、店情報、商品情報、他のレビュー) へのアクセスを可能とするリンクを提示した。これは使用したデータセットの元となっている楽天市場におけるレビューの提示方法とほぼ同様である。これをレビュー 30 件分繰り返した。

次に、1 週間後、同じ被験者に対して、提案システムを利用したうえでレビューが信頼できるかどうかのアンケートに答えてもらった。提示したレビューは、システムを利用しない実験に用いたのと同じレビュー 30 件であるが、提示する順番をランダムに変更した。このとき、各指標のスコア、各指標に関する情報へのアクセスを可能とするリンクに加えて、システムを利用しない場合に提示したのと同様のレビュー内容や各種情報 (投稿者履歴、店情報、商品情報、他のレビュー) へのアクセスを可能とするリンクも提示した。

そして、提案システムを利用しない場合と利用する場合の被験者の信頼性に対する判断の違いを比較した。

### 5.2.1 提示するレビュー

被験者に提示するレビューは、1 つの指標でスコアの高い (スコアが 4 以上) レビュー 12 件、2 つの指標においてスコアの高いレビュー 6 件、3 つの指標 (類似性・集中性・情報性の組合せ) でスコアの高いレビュー 1 件、どの指標においてもスコアの高くない (スコアが 4 未満) レビュー 11 件の合計 30 件である。提示したレビューの内訳を表 7 に示す。また、30 件の内 15 件が投稿者履歴を見ることが

表 7 提示するレビューの内訳 (件数)

Table 7 The number of reviews in the questionnaire.

指標	件数
類似性	7
協調性	2
集中性	3
情報性	0
類似性+協調性	3
類似性+集中性	2
類似性+情報性	1
協調性+集中性	0
協調性+情報性	0
集中性+情報性	0
類似性+集中性+情報性	1
スコアの高くはないもの	11
計	30

可能なレビューで、残り 15 件は投稿者履歴を見ることができないレビューである。

### 5.2.2 アンケート

#### 信頼性に関するアンケート (システムを利用しない場合)

提案システムを利用しない場合の信頼性に関するアンケートの質問項目は以下のとおりである。これは、レビュー 1 件ずつ、30 件すべてに対して行う。

(1) レビューが信頼できるかどうか判定してください

- (a) 信頼できる
- (b) まあまあ信頼できる
- (c) どちらともいえない
- (d) あまり信頼できない
- (e) 信頼できない

(2) 何を理由に判断しましたか? (複数回答可)

- (a) レビュー本文
- (b) ユーザの履歴
- (c) 商品のレビュー
- (d) ショップの履歴
- (e) その他

このとき、「何を理由に判断しましたか?」の回答の“ユーザの履歴”は、レビューを投稿した投稿者の履歴を表し、“商品のレビュー”は、同じ商品に対して他のユーザが投稿したレビューを表す。そして、“ショップの履歴”は、店その他の商品情報や他の商品のレビューを表す。これらの説明をあらかじめ行ったうえで、被験者には実験を始めてもらった。

#### 信頼性に関するアンケート (システムを利用する場合)

提案システムを利用する場合の信頼性に関するアンケートの質問項目は提案システムを利用しない場合の質問項目と同様であるが、(2)の質問項目の選択肢を以下のように変更する。これは、レビュー 1 件ずつ、30 件すべてに対して行う。

(2) 何を理由に判断しましたか? (複数回答可)

- (a) レビュー本文
- (b) ユーザの履歴
- (c) 商品のレビュー
- (d) ショップの履歴
- (e) 類似性スコア
- (f) 類似性リンク
- (g) 集中性スコア
- (h) 集中性リンク
- (i) 協調性スコア
- (j) 協調性リンク
- (k) 情報性スコア
- (l) その他

さらに、実験の終わりに下記の質問について回答してもらった。

- システムを使ったことで信頼性を判断する必要性を感じましたか?

- (1) はい
- (2) いいえ
- (3) どちらともいえない

以上が、本研究における実験で行ったアンケート内容である。

### 5.3 実験結果と考察

#### 5.3.1 アンケート結果

表 8、表 9 に信頼性に関するアンケート結果を示す。表 8 は、信頼性に関するアンケートの回答結果である。行は各質問項目に、列は各レビューに対応する。このとき、投稿者履歴の欄は、値が 0 であれば、そのレビューを投稿した投稿者の履歴を見ることができないレビューであることを表し、1 であれば投稿者の履歴を見ることができるレビューであることを表す。そして、各セルの値は、各レビューに対する各回答を行った被験者の人数である。投稿者履歴を見ることができないレビューでは、ユーザ履歴や協調性スコア・リンクを判断材料とすることができないため、“-”で表す。また、①は提案システムを利用しない場合の回答結果であり、②は提案システムを利用する場合の回答結果である。このとき、「信頼できる」・「まあまあ信頼できる」は、“信頼できる”とし、「あまり信頼できない」・「信頼できない」は、“信頼できない”として扱う。また、表 9 は、実験の最後に行った、「システムを使ったことで信頼性を判断する必要性を感じましたか?」に対する回答結果である。

まず、表 8 の「レビューが信頼できるかどうか判定してください」の項目に対する被験者間の回答 (信頼できる・どちらともいえない・信頼できない) の一致度を調査した。一致度を求めるための評価尺度には、Fleiss の  $\kappa$  係数 [20] を用いる。これは、値が 1 であれば、被験者間の回答の完

表 8 信頼性に関するアンケート結果の内訳 1  
Table 8 The result of questionnaire 1.

レビュー ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
投稿者履歴	1	1	0	0	0	0	1	1	1	0	1	0	0	1	0	0	1	1	1	1	0	0	1	0	0	1	0	1	0	1	
①	レビューが信頼できるかどうか判定してください																														計
信頼できる	17	15	7	4	1	20	20	9	11	6	15	6	10	13	3	0	13	19	2	15	13	18	16	8	16	7	13	1	1	12	311
どちらともいえない	1	4	8	2	5	0	0	8	3	7	2	12	5	5	11	8	4	1	2	2	5	1	2	11	2	6	6	7	4	2	136
信頼できない	3	2	6	15	15	1	1	4	7	8	4	3	6	3	7	13	4	1	17	4	3	2	3	2	3	8	2	13	16	7	183
	何を理由に判断しましたか？																														
レビュー本文	20	16	20	19	19	21	20	18	16	19	17	15	15	15	18	19	14	17	18	19	17	20	16	19	21	17	20	20	20	17	542
ユーザの履歴	8	12	-	-	-	-	10	7	9	-	9	-	-	1	-	-	10	10	9	-	-	8	-	-	12	-	9	-	9	133	
商品のレビュー	4	10	8	7	9	8	7	8	10	8	7	9	11	13	9	8	11	8	4	6	11	3	5	10	7	3	9	4	4	9	230
ショップの履歴	0	0	1	1	2	0	1	5	1	1	1	1	3	1	2	2	3	1	0	1	1	1	0	3	0	0	0	1	2	3	38
その他	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
②	レビューが信頼できるかどうか判定してください																														計
信頼できる	19	21	3	6	0	11	21	12	3	4	4	1	3	10	2	0	3	11	0	18	4	21	19	0	19	7	19	1	1	3	246
どちらともいえない	1	0	2	6	1	4	0	8	2	5	8	1	3	2	4	1	3	8	0	2	4	0	1	3	1	5	2	2	5	5	89
信頼できない	1	0	16	9	20	6	0	1	16	12	9	19	15	9	15	20	15	2	21	1	13	0	1	18	1	9	0	18	15	13	295
	何を理由に判断しましたか？																														
レビュー本文	21	20	14	16	19	19	20	18	8	15	17	14	11	13	16	18	8	17	15	20	14	20	18	8	20	13	19	19	21	11	482
ユーザの履歴	4	6	-	-	-	-	8	4	2	-	4	-	-	4	-	-	1	6	4	3	-	-	6	-	-	4	-	3	-	6	65
商品のレビュー	3	5	4	2	2	2	4	3	4	4	4	4	2	5	2	4	2	1	2	3	5	2	3	2	2	3	5	1	2	1	88
ショップの履歴	1	0	1	0	0	0	0	0	2	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
類似性スコア	12	11	14	7	17	11	15	9	0	12	13	16	1	6	18	16	1	4	16	14	10	12	14	17	10	12	12	15	8	12	335
類似性リンク	5	1	9	2	14	14	8	3	0	14	12	20	1	1	17	12	0	3	17	5	9	2	2	16	2	12	3	14	8	13	239
集中性スコア	7	6	6	5	17	0	6	4	4	8	1	1	17	9	10	12	5	2	0	7	13	3	6	9	5	3	5	3	0	0	174
集中性リンク	0	0	5	5	9	0	1	0	2	8	0	1	13	8	9	9	7	0	0	0	13	1	1	9	1	0	0	3	0	1	106
協調性スコア	7	5	-	-	-	-	6	5	15	-	6	-	-	1	-	-	15	12	0	8	-	-	7	-	-	6	-	1	-	9	103
協調性リンク	0	0	-	-	-	-	1	0	11	-	5	-	-	0	-	-	14	10	0	0	-	-	0	-	-	7	-	0	-	9	57
情報性スコア	7	2	1	2	10	1	8	2	2	3	0	0	1	1	1	3	2	1	0	7	2	4	2	0	4	1	5	0	2	1	75
その他	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

表 9 信頼性に関するアンケート結果の内訳 2  
Table 9 The result of questionnaire 2.

	システムを使ったことで信頼性を判断する必要を感じましたか？
はい	20 (95%)
いいえ	0 (0%)
どちらともいえない	1 (5%)

表 10 回答の一致度

Table 10 The degree of coincidence of answers.

	一致度 (κ)
提案システムを利用しない場合	0.22
提案システムを利用する場合	0.41

全な一致を表し、0 以下であれば被験者間の回答が一致していないことを表す係数である。表 10 がその結果である。提案システムを利用する場合の被験者間の回答の一致度は、利用しない場合に比べて、増加していることが分かる。

また、表 8 の「レビューが信頼できるかどうか判定してください」の項目に関する回答の合計数に注目する。提案システムを利用しない場合に比べて、提案システムを利用する場合では、「どちらともいえない」という回答件数が

減っていることが分かる。つまり、各指標に関するスコアやリンク先の情報により、被験者は「信頼できる」か「信頼できない」かの判断を行うことができるようになったことが確認できる。

これらの結果から、提案システムによって提示される情報は、信頼性を判断するための良い判断材料となっていると考えられる。つまり、提案システムによって信頼性を判断することが容易になり、有効な信頼性の判断支援をすることが可能と考えられる。

次に、表 8 の「何を理由に判断しましたか？」の項目に関する回答に注目する。これは、各レビューに対して、被験者がどの情報を参考にして、信頼できるかどうかを判断したかを表す。

まず、提案システムを利用しない場合について考察する。「レビュー本文」を参考にした被験者は多く、1つのレビューあたり平均 18 人が参考にしていて、また、「ユーザの履歴」は、投稿者履歴を見ることができるレビュー (15 件) に限り、1つのレビューあたり平均 9 人が参考にしており、「商品のレビュー」は、1つのレビューあたり平均 8 人が参考にしていて、「ショップの履歴」は、参考にされる機会がほとんどなかった。これは、事前実験 (表 5, 表 6)

より、スパムをあまり意識しない人が多く、“レビュー本文”以外を参考にする被験者が少なかったという結果と一致している。

次に、提案システムを利用する場合について考察する。1つのレビューあたり“レビュー本文”は平均16人、“ユーザの履歴”は投稿者履歴を見ることができるレビューで平均4人、“商品のレビュー”は平均3人が参考にしており、“ショップの履歴”はほとんど参考にされなかった。提案システムを利用しない場合に比べて、これらの情報を参考にする被験者は減っていることが分かる。これは、各指標に関するスコアやリンクを重視して判断する被験者の存在によるものである。特に、“類似性スコア”・“類似性リンク”・“集中性スコア”・“集中性リンク”・“協調性スコア”に関しては、提案システムを利用する場合における“ユーザの履歴”や“商品のレビュー”よりも参考にされる機会が多かった。このことから、提案システムにより、被験者の信頼性に対する判断の支援を行えていることが確認できる。

最後に、「システムを使ったことで信頼性を判断する必要性を感じましたか？」に対する回答(表9)について考察する。事前実験(表5)より、被験者はふだんはスパムをあまり意識しないことが確認されている。しかしながら、「システムを使ったことで信頼性を判断する必要性を感じましたか？」に対する回答では20人(95%)の被験者が“はい”と回答し、信頼性に関する意識の向上が見られた。つまり、提案システムを利用することで、ユーザに対して信頼性に対する意識を促すことが可能と考えられる。以上より、提案システムを利用することで、ユーザに対して、信頼性への意識を促すとともに、有効な信頼性の判断支援を行うことができる。

5.3.2 考察

アンケート結果から、提案システムを利用することで、有効な判断支援が行えることが分かった。ここからは、提

案システムを利用しない場合と利用する場合の、レビューに対する判断の変化について、さらなる考察を行う。

そのため、表8の回答結果から、レビューを5つのタイプに分類した。表11に、レビューの分類と各指標に関するスコア値(0以上5以下)、そして信頼できるか否かの回答結果を示す。以降では、表11をもとに、各タイプの説明と考察を行っていく。

タイプ1：提案システムを利用しない場合では、被験者の判断がわかれている(どの回答も過半数を超えていない、もしくは“どちらともいえない”の回答が過半数を超えている)が、提案システムを利用する場合では、被験者の判断がまとまっている(“信頼できる”、もしくは“信頼できない”が過半数を超えている)レビュー。

表11より、レビューID3, 8, 10, 12, 13, 15, 24がタイプ1に該当する。

考察：ID8とID13のレビューを除くこのタイプのすべてのレビューは、類似性スコアは5であり、複製されたレビューである。また、投稿者履歴を見ることができない。そのため、提案システムを利用しない場合では、信頼性を判断するための情報が少なく、主に“レビュー本文”から判断を行った結果、被験者の判断はわかれた(もしくは判断がつかない)と考えられる。しかしながら、提案システムを利用する場合では、類似性スコアを示すとともに、類似性リンクにより、レビューが複製されたものであることが確認できる。そのため、“信頼できない”として、被験者の判断がまとまったと考えられる。

また、ID13のレビューに関しては、集中性スコアが高かった。そのため、システムを利用した場合には“信頼できない”とした被験者が増えたと思われる。

ID8のレビューは、提案システムを利用する場合、どの指標に関するスコアも高くなかったため、それを理由に“信頼できる”とした被験者が増えたと思われる。

表 11 レビューのスコア値と回答結果(タイプ別)

Table 11 The classification of answers and scores in the questionnaire.

タイプ	1								2				3						4										その他			
レビュー ID	3	8	10	12	13	15	24	4	11	14	9	17	21	30	1	2	5	7	16	19	20	22	23	25	27	28	6	18	26	29		
投稿者履歴	0	1	0	0	0	0	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	0	1	0	0	1	0	1	1	0		
類似性スコア値	5	1.5	5	5	1.8	5	5	0.6	5	1	0.4	0.9	3.2	5	0.4	0.7	5	0.5	5	5	0.6	0.4	0.4	0.3	0.9	5	5	0.7	5	5		
協調性スコア値	0	0	0	0	0	0	0	4.7	0	5	4.9	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.7	0	
集中性スコア値	3.1	0	3	0	5	3.6	4.9	2	0	3	3	3.3	4.2	0	0	0	4.4	0	4.4	0	0	0	0	0	0	0	0	4.7	0	0		
情報性スコア値	2.4	2.6	3	2.7	3.2	2.5	3	2.8	2.9	2.6	2.6	2.5	2.7	2.8	1.1	1.9	5	1.3	3.3	2	1.5	1.2	2.6	0.4	2.1	3	0.9	2.5	2.4	5		
①																																
信頼できる	7	9	6	6	10	3	8	4	15	13	11	13	13	12	17	15	1	20	0	2	15	18	16	16	13	1	20	19	7	1		
どちらともいえない	8	8	7	12	5	11	11	2	2	5	3	4	5	2	1	4	5	0	8	2	2	1	2	2	6	7	0	1	6	4		
信頼できない	6	4	8	3	6	7	2	15	4	3	7	4	3	7	3	2	15	1	13	17	4	2	3	3	2	13	1	1	8	16		
②																																
信頼できる	3	12	4	1	3	2	0	6	4	10	3	3	4	3	19	21	0	21	0	0	18	21	19	19	19	1	11	11	7	1		
どちらともいえない	2	8	5	1	3	4	3	6	8	2	2	3	4	5	1	0	1	0	1	0	2	0	1	1	2	2	4	8	5	5		
信頼できない	16	1	12	19	15	15	18	9	9	9	16	15	13	13	1	0	20	0	20	21	1	0	1	1	0	18	6	2	9	15		

**タイプ2:** 提案システムを利用しない場合では、被験者の判断がまとまっている(“信頼できる”, もしくは“信頼できない”が過半数を超えている)が、提案システムを利用する場合では、被験者の判断がバラついている(どの回答も過半数を超えていない)レビュー。

表 11 より, レビュー ID4, 11, 14 がタイプ2 に該当する。

**考察:** 提案システムを利用しない場合, これらのレビューは, “信頼できる”, もしくは“信頼できない”に判断がまとまっていた。しかし, これらのうち, 被験者の多くが“信頼できない”としたレビューは, 提案システムにおいて特にスコアの高い指標はなく, 逆に, 被験者の多くが“信頼できる”としたレビューのスコアに関しては, 何かしらの指標のスコアが高かった。そのため, 提案システムを利用する場合に, 判断がバラついたりと考えられる。これにより, 被験者たちは, 単にスコアのみを鵜呑みにしたのではなく, スコアを参考に主体的に判断を行えていることがうかがえる。

**タイプ3:** 提案システムを利用しない場合と提案システムを利用する場合とでは, 被験者の判断が逆転しているレビュー (“信頼できる”が過半数を超えていたのに, 提案システムだと“信頼できない”が過半数を超えている。または, その逆)。

表 11 より, レビュー ID9, 17, 21, 30 がタイプ3 に該当する。

**考察:** これらのレビューは, 協調性スコア, もしくは集中性スコアが高い。しかし, 提案システムを利用しない場合, 多くの被験者は“信頼できる”と判断している。これは, 被験者がレビューの投稿時間を意識していないこと(事前実験の表 6 参照)や, 他の投稿者と協調しているかどうかの判断材料にたどり着いていないことが理由として考えられる。提案システムでは, 協調性や集中性に関するスコアや情報を提示することで, 被験者に対して怪しさを気づかせることができる。そのため, 被験者の判断は“信頼できない”に変わったと考えられる。

**タイプ4:** 提案システムを利用しない場合でも, ある程度被験者の判断がまとまっているが (“信頼できる”, または“信頼できない”が過半数を超えている), 提案システムを利用する場合では, より一層, その被験者の判断がまとまっているレビュー。

表 11 より, レビュー ID1, 2, 5, 7, 16, 19, 20, 22, 23, 25, 27, 28 がタイプ4 に該当する。

**考察:** これらのレビューは, 投稿者の履歴を見ることができるとも多く(12件中7件), また“レビュー本文”から, 信頼できるかどうか判断がしやすいものであった(たとえば, 意味の分からない文章のレビューは“信頼できない”, 逆にしっかりと細かいところまでかかれ

た文章のレビューは“信頼できる”, など)。そのため, 提案システムを利用しない場合でも, 被験者の回答はまとまっていた。また, これらのレビューのうち, 被験者の多くが“信頼できない”としたレビューは, 提案システムにおいても何かしらの指標のスコアが高く, 逆に, 被験者の多くが“信頼できる”としたレビューのスコアに関しては, 特にスコアの高い指標はなかった。そのため, 提案システムを利用することで, 被験者の判断がより確かなものとなり, その結果, 回答がより一層まとまったと考えられる。

**その他:** タイプ1~タイプ4のどれにも属さないレビューであり, レビュー ID6, 18, 26, 29 がタイプ5 に該当する。

レビュー ID6 は商品に対する情報が詳細に書かれたレビューであり, ID18 は投稿者の履歴から信頼できそうなレビューを多く投稿している投稿者のレビューであった。そのため, 被験者は“信頼できる”と判断していた。しかし, 提案システムを利用する場合, スコアが高い指標があることが確認できる。そのため, “信頼できない”, “どちらともいえない”の回答が増えたが, “信頼できる”と回答した被験者が多いのは変わらなかった。

レビュー ID26 は, コミックに関するレビューであり, シリーズものである。そのため, このレビューの投稿者は, 同じシリーズの異なる巻に, すべて同じ文章を投稿していた。このような傾向は, 本ジャンルには多く見られ, 被験者もそれをスパムとして怪しいとするかどうかの判断が分かれた。そして, 提案システムを利用する場合においても, その傾向は変わらなかった。

レビュー ID29 は, レビューの文章があきらかにおかしかったため(「。。。。。。。」という句点のみのレビューであった), 提案システムに関係なく, 多くの被験者は“信頼できない”と判断した。

これらの結果からも, どの指標を重視するかや, 提示される根拠に基づいてどのように判断するかはユーザによって異なるが, 各ユーザは主体的に判断していることがうかがえる。

以上より, レビューの分類を行い, 各タイプごとに考察を行った。この結果から, 特にタイプ1 やタイプ3 に見受けられるように, ふだん被験者が気づくことのできない情報に関して, 提案システムは提示することができていることが分かる。また, そのようなレビューに対して, 特に提案システムは有効だと考えられる。

## 6. おわりに

本研究では, スпам検知に関する既存研究から, ec サイトにおけるレビューの信頼性に関する4つの指標を新たに

定義した。加えて、各指標を用いることで、レビューの信頼性の判断を支援するためのシステムを提案した。また、提案システムの有効性を検証するために、提案システムを利用しない場合と利用する場合とで、レビューが信頼できるかどうかのアンケートを行った。この結果、被験者に対して、信頼性の意識を促すことができることを確認した。また、提案システムを利用しない場合とする場合との、被験者間の信頼性の判断の一致度を比較した。その結果、提案システムを利用する場合の方が、利用しない場合と比べて、被験者間の判断の一致度が増加することが確認できた。これは、提案した指標が、被験者に対して有用な情報になり、信頼性を判断することが容易になったことを意味する。つまり、被験者に対して、提案システムは有効な判断支援が行えることが分かった。

今後の課題として、より大規模なデータセットを利用することがあげられる。今回は、楽天データの「みんなのレビュー・口コミデータ」の2010年1月1日から1月31日の1カ月間のみのレビューデータを用いた。しかしながら、サクラは長期間的にスパムを投稿することも考えられる。文献[9]では、1年間のデータセットを用いており、長期的にレビュースパムを投稿するサクラの存在も確認されている。そのため、長期的なデータセットをもとに、各指標のスコアを求めた場合の提案システムの有効性を調査する必要がある。また、今回は店を対象としたレビューと商品を対象としたレビューの区別をしていないが、今後は、その区別をすることで各指標の傾向に関しても考察する必要がある。これには、本文の内容によるレビューの分類が必要となる。

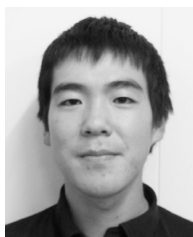
さらに今回のシステムでは、4つの既存のスパム検知手法で用いられているスパムのよく知られた特徴を指標化して提示したが、スパムをスパムとして判断するのに、この4つの指標で十分というわけではない。また、スパム検知手法としては自然言語処理を用いた様々な手法も新たに提案されている[21], [22]。これらは、よりレビュー本文の意味を分析し、評価値と本文の整合性や使われている単語の品詞や文脈などに着目した手法である。今後、これらの検知手法における観点についても指標化し、システムに導入する可能性も検討していく。

謝辞 本研究を進めるにあたり、データを提供していただきました楽天技術研究所に心より御礼申し上げます。

## 参考文献

- [1] Huang, S., Shen, D., Feng, W., Zhang, Y. and Baudin, C.: Discovering clues for review quality from author's behaviors on e-commerce sites, *Proc. 11th International Conference on Electronic Commerce*, pp.133-141 (2009).
- [2] Chevalier, J. and Mayzlin, D.: The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research*, Vol.43, No.3, pp.345-354 (2006).
- [3] Neylon, S.B.T., Hannan, K., Reis, G.A., McDonald, R. and Reynar, J.: Building a Sentiment Summarizer for Local Service Reviews, *Proc. WWW Workshop on NLP in the Information Explosion Era* (2008).
- [4] Pang, B., Lee, L. and Vaithyanathan, S.: Thumbs up? Sentiment Classification using Machine Learning Techniques, *Proc. Conference on Empirical Methods in Natural Language Processing*, pp.79-86 (2002).
- [5] Liu, J., Cao, Y., Lin, C.Y., Huang, Y. and Zhou, M.: Low-Quality Product Review Detection in Opinion Summarization, *Proc. 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp.334-342 (2007).
- [6] Hong, Y., Lu, J., Yao, J., Zhu, Q. and Zhou, G.: What Reviews are Satisfactory: Novel Features for Automatic Helpfulness Voting, *Proc. 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.495-504 (2012).
- [7] 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).
- [8] Jindai, N. and Liu, B.: Opinion Spam and Analysis, *Proc. International Conference on Web Search and Web Data Mining*, pp.219-230 (2008).
- [9] Xie, S., Wang, G., Lin, S. and Yu, P.S.: Review Spam Detection via Temporal Pattern Discovery, *Proc. 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.823-831 (2012).
- [10] Ott, M., Choi, Y., Cardie, C. and Hancock, J.T.: Finding Deceptive Opinion Spam by Any Stretch of the Imagination, *Proc. 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp.309-319 (2011).
- [11] 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).
- [12] Yamamoto, Y. and Tanaka, K.: Enhancing Credibility Judgment of Web Search Results, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pp.1235-1244 (2011).
- [13] Metzger, M.J., Flanagin, A.J. and Zwarun, L.: College student web use, perceptions of information credibility, and verification behavior, *Computer and Education*, Vol.41, No.3, pp.271-290 (2003).
- [14] Broder, A.Z.: On the Resemblance and Containment of Documents, *Proc. Compression and Complexity of Sequences*, p.21 (1997).
- [15] Zaki, M.J., Parthasarathy, S., Ogihara, M. and Li, W.: New Algorithms for Fast Discovery of Association Rules, *Proc. 3rd International Conference on Knowledge Discovery and Data Mining*, pp.283-286 (1997).
- [16] Zhang, S., Chakrabarti, A., Ford, J. and Makedon, F.: Attack Detection in Time Series for Recommendation Systems, *Proc. 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.809-814 (2006).
- [17] Kleinberg, J.: Bursty and hierarchical structure in streams, *Proc. 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.91-101 (2002).
- [18] Johnson, M.K. and Raye, C.L.: Reality monitoring, *Psychological Review*, Vol.88, No.1, pp.67-85 (1981).

- [19] Rayson, P., Wilson, A. and Leech, G.: Grammatical word class variation within the British National Corpus sampler, *Language and Computers*, Vol.36, No.1, pp.295–306 (2002).
- [20] Fleiss, J.L.: Measuring nominal scale agreement among many raters, *Psychological*, Vol.76, No.5, pp.378–382 (1971).
- [21] Sharma, K. and Lin, K.: Review spam detector with rating consistency check, *Proc. 51st ACM Southeast Conference*, No.34 (2013).
- [22] Morales, A., Sun, H. and Yan, X.: Synthetic review spamming and defense, *Proc. 22nd International Conference on World Wide Web Companion*, pp.155–156 (2013).



伊木 惇

2011年広島大学工学部第二類（電気系）卒業。2014年同大学大学院工学研究科博士課程前期修了。同年中国電力株式会社入社。



亀井 清華（正会員）

2001年広島大学工学部第二類（電気系）卒業。2003年同大学大学院工学研究科博士課程前期修了。2006年同博士課程後期修了。博士（工学）。2006年鳥取環境大学助手。2008年広島大学助教。2012年同大学准教授。

分散アルゴリズムの研究に従事。電子情報通信学会, IEEE, IEEE CS, ACM 各会員。



藤田 聡（正会員）

1985年広島大学工学部第二類（電気系）卒業。1990年同大学大学院工学研究科博士課程後期修了。工学博士。現在、広島大学大学院工学研究院教授。並列・分散アルゴリズムやネットワークアルゴリズム, グラフアルゴリズム

等の研究に従事。電子情報通信学会, 日本応用数理学会, IEEE CS, SIAM 各会員。