

中澤ゼミ

## 信頼性の高い口コミ情報検索掲示板

a1200318 鈴木香織

### 1 はじめに

現在、消費者は商品購入やサービスを利用する場合に、企業側による情報よりも利用者の生の声である「口コミ情報」を重視する傾向にある[15]。なぜなら、企業が提供する広告やカタログ情報などは、「良いことが書いてあって当たり前」であって、マイナス面がわかりにくいからである。

Amazon.co.jp や価格.com の大規模な口コミ掲示板が成り立っているのは、消費者がメーカー側の情報だけでは満足できず、より多角的な情報を求めているということを表している[16]。一方で口コミ情報の中には、辛口なコメントも見受けられる場合がある。そのため、消費者の視点から率直な意見が述べられているのでプラス面だけでなくマイナス面も知ることができる。

消費者に口コミ情報を提供するものとしてインターネット上では、個人のホームページ、電子掲示板、メーリングリスト、メールマガジンなどが挙げられる。これらの中で、個人のホームページ、メーリングリスト、メールマガジンは、情報が一方的に提供されるため、消費者同士が意見交換や情報共有をすることが難しい。一方、電子掲示板は意見交換や情報共有に双方向性があるため、口コミ情報を提供するものとして適している。

本研究では、本当に信頼性のある情報かどうか判断が難しい点、膨大な情報の中から利用者の望む情報を検索しにくいといった電子掲示板の問題点を改善するために、電子掲示板に情報推薦システムを組み合わせ、消費者が信頼性のある情報を得やすい検索環境を支援するシステムを提案する。

### 2 電子掲示板

電子掲示板(以下、掲示板と表記する。)とは、ブラウザを通じてウェブページに文章を残し、他の人と情報共有や意見交換をするための仕組みである。

#### 2.1 電子掲示板の運用形態

掲示板の運用形態として会員制掲示板と匿名掲示板が挙げられる。会員制掲示板では書き込みや閲覧の際に登録が必要となる。このメリットとして、なりすまじや匿名の書き込みをある程度防ぐことができるため、信頼性のある情報を得やすい。さらに、登録の際にユーザのプロファイル<sup>1</sup>情報を集めることができるため、ユーザが必要とする情報だけを提供することのできる、新しいシステムを構築することも可能である。

一方、匿名掲示板では書き込みや閲覧の際に登録を必要としない。そのため、匿名で気兼ねなく意見を投稿することができ、情報が集まりやすいというメリットが挙げられる。しかし、匿名であるが故に無責任で感情的なやりとりから、フレーミング<sup>2</sup>と呼ばれる敵意的な行動が誘発されやすいという問題点が挙げられる。

#### 2.2 口コミ掲示板の信頼性

本研究では、口コミ掲示板を実際に商品・サービスを利用、購入した消費者の意見や感想といった情報を共有するための掲示板と定義する。

一般的に、新しく信頼性の高い情報ほど価値が高いと言われる[14]。しかし、既存の口コミ掲示板では、新しい知識が出現しているのにも関わらず、古い知識を探し当ててしまうことがある。そのため、知識の時間に対する価値の変化を考慮する必要があり、情報の新しい記事を取捨選択することができるシステムが必要となる。

また、口コミ掲示板では、ユーザー一人一人の趣味・嗜好に合った情報ほど価値は高くなる傾向にある。そのため、他のユーザとどの程度特性が似ているかを判断し、フィルタリングする必要がある。

この他に、匿名性の高い口コミ掲示板では発言の責任が追及されないため、フレーミングの誘発等により情報の信頼性が低下する。そのため、フレーミングを軽減し、レビュアーや記事の中身を取捨選択をする事ができるシステムが求められる。

### 3 情報推薦システム

情報推薦システムとは、情報収集支援アプローチの一つである。これは、ユーザに対して有益な情報を推薦する手法である。従来の情報検索システムでは、利用者の要求が曖昧な場合必ずしも利用者の望む情報が検索結果に反映されない場合がある。それを防ぐためにシステム側が利用者のプロフィールや過去の履歴、ユーザによるレーティング情報などを元に利用者へ情報を推薦する仕組みを情報推薦システムという[14]。情報推薦システムの応用例として、Amazon.co.jp では消費者の購入履歴データから消費者の嗜好に合った商品を推薦するシステムが挙げられる。

本研究では、電子掲示板に情報推薦システムを取り入れ、ユーザが望む信頼性の高い情報を得やすい環境を構築する。

#### 3.1 レーティング(情報の格付け・評価)

情報推薦システムに用いられる技術の一つにレイティ

<sup>1</sup> ユーザの設定やデータを履歴として登録したもの。この場合、ユーザの行動履歴を指す。

中澤ゼミ

ングが挙げられる。レイトイングとは、情報に対して一定の基準に沿ったラベル付け(格付け)を行うことである。発信者自身が情報にレイトイングを行う、「セルフ・レイトイング」と第三者が付加的に情報のレイトイングを行う「サードパーティ・レイトイング」がある。レイトイングした結果もユーザ行動履歴と同様に、フィルタリングルールで処理される。

3.2 情報フィルタリング

情報推薦システムに用いられるもう一つの技術が情報フィルタリングである[14]。これは、サーバ側がユーザの設定したプロフィール情報のレベルや目的に合わせて、ユーザが必要とする情報のみを提供することを目的としている。そのために、ユーザに送る情報を除外する方法をフィルタリングルールという。このシステムによって、ユーザは情報を効率的に得ることができる。

4 信頼性の高い口コミ情報検索掲示板

4.1 ユーザ行動履歴とレイトイングを応用

信頼性の高い口コミ情報を効率的に検索できるようにするため、ユーザ行動履歴とレイトイングを応用した掲示板システムを提案する。従来、Amazon.co.jp<sup>3</sup>などで使われているユーザ行動履歴を使用したシステムは、クッキー<sup>4</sup>にユーザ行動履歴を保存して情報推薦を行っている。しかし、クッキーに保存される情報は一時的なもので、ユーザ毎に会員情報やレイトイングのための各種評価値、評価値による算出値を格納することができない。そのため、データベースを用いたシステムが必要である。データベースを用いたシステムには、行動履歴や記事に対するレイトイング情報、フィルタリングルールが格納される。このフィルタリングのルールをユーザ毎に適用することによって、ユーザが大量の記事から信頼性のある記事のみを取捨選択支援することができる。このため、匿名掲示板ではなく会員制掲示板の運営形態をとる。

4.2 システム概要

ユーザ行動履歴とレイトイングを応用する口コミ情報検索掲示板の構成要素として、第一に、会員情報やレイトイングのための各種評価値、評価値による算出値を格納するためにデータベース、第二に、データベースに格納した値を提案するフィルタリングルールに基づき、記事の取捨選択を支援するフィルタリングシステムが挙げられる。このような二つの要素でシステムを構成する。

4.3 会員登録の流れ

4.3.1 会員登録

会員登録は、メールアドレスとIDで登録することができるという簡易的な方式でよい。登録を簡易にすることで、情報が集まりにくい問題を軽減する。任意でプロフィール情報をデータベースに登録することで、ユーザ専用のフィルタリング環境を提供することができる。また、会員登録によってなりすましや匿名の書き込みを防ぐことができるためフレーミング行為の抑制につながる。

4.3.2 登録手順

フィルタリングルールを用いて記事の取捨選択を行いたいユーザだけが性別や年代、特徴や好みといった任意のユーザ情報を入力し会員登録をする。情報を多く入力するほど精度の高いフィルタリングをすることができる。入力された情報はデータベースに格納され、フィルタリングの際に利用される。フィルタリングルールを使用しないユーザは、記事を取捨選択せず単なる閲覧だけを行うことができる。

4.4 システムの手順の流れ

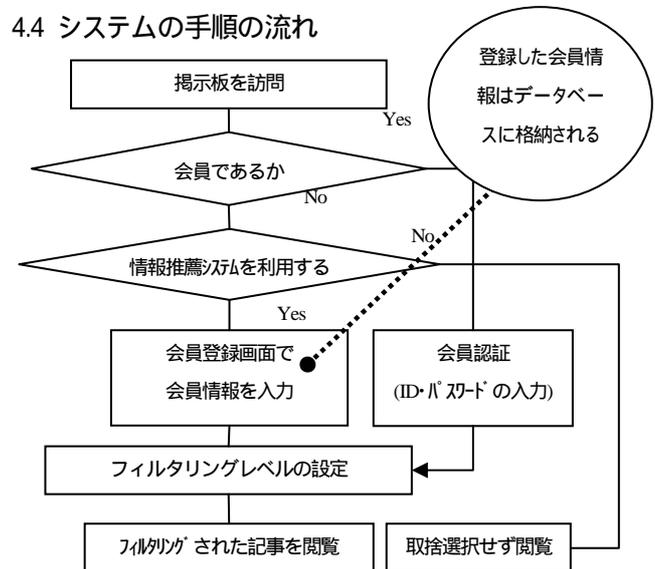


図1: 掲示板システムの利用手順

4.5 フィルタリングの計算方法

掲示板に表示されている記事をクリックすると図2のように記事の内容が最初に表示される。

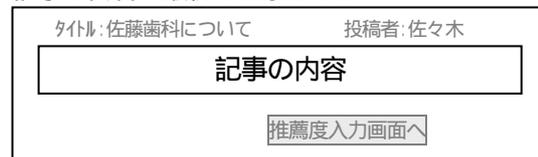


図2: 記事の評価をしていない場合の記事画面

4.5.1 推薦度

推薦度とはレビュアーが記事の内容を評価した点数であり、会員が評価入力画面で入力する。図2の記事を読んだレビュアーが推薦度入力ボタンから図3の推薦度入力画面へジャンプする。そこで、1点から4点の点数を選択し、記事の内容について評価をする。4段階評価としているのは、中心化傾向を避けるためである。

<sup>2</sup>敵意的な言語行動を指す。

<sup>3</sup>インターネットを利用した本などの通信販売サイトである。

<sup>4</sup>Webサイトの提供者が、Webブラウザを通じて訪問者のコンピュータに一時的にデータを書き込んで保存させる仕組み

中澤ゼミ

図3:推薦度入力画面(記事の内容の評価)

4.5.2 信頼度

信頼度とは、記事に対して評価をしているレビュアー(自分を含む)に対する信頼の度合いである。信頼度は1から5の範囲の値をとり、初期値を3とする。これは未知のレビュアーに対する信頼度を中間値である3に設定するためである。また、自分自身に対する信頼度は5点とする。レビューを閲覧する度にレビュアーに信頼度を評価するので、信頼度合計を閲覧回数で割った平均信頼度で表される(以後、平均信頼度を単に信頼度と呼ぶことにする)。

図4:推薦度出力・信頼度入力画面(レビュアーの評価)

図3と図4の評価入力画面で入力した推薦度、信頼度、会員登録の際に入力した消費者情報の3つの情報に基づいて記事の取捨選択を行う。情報推薦システムを利用しない人は、フィルタリングルールの利用はできず、すべての記事が表示される。

4.5.3 加重平均推薦度

加重平均推薦度とは信頼度と推薦度によって算出された記事一つ一つに対する総合的な推薦度である。この値が高いほど信頼性が高い。単純平均ではなく、評価したレビュアーの信頼度合計に占める各信頼度の割合を重みとして、加重平均値を求めている。

表1:会員鈴木に合った記事を取捨選択する場合

	五十嵐	大竹	佐々木	香織	智子	馬場	沼田	宮島
	5	1	2	5	5	4	3	5
記事A	2		4		2	3		
記事B	4			1				
記事C					3			
記事D	3		3	3		3		
記事E								
記事F	3			3			3	3
記事G		2			3		4	

加重平均推薦度は次式で求めることができる。

$$U_i \times \frac{V_i}{V_i}$$

ここで  $U_i$  は  $i$  番目のレビュアーに対する信頼度を、 $V_i$  はその記事に対する  $i$  番目のレビュアーが評価した推薦度を表すものとする。例外として、評価をしたレビュアー全てが同じ推薦度の場合、算出値が同じになってしまうことが挙げられる。そのため、算出値が同じになってしまった場合は、レビュアーの信頼度を比べ、高い順番に記事を表示させる。また付加情報は、レビュアーの人数を表示するというものである。

4.5.4 類似度(ユーザ行動履歴を用いた記事推薦)

類似度とは、会員登録の際に入力した任意の会員情報を他の会員と比較し、どの程度特性が似ているかを表したものである。類似度が高い程、対象のユーザと比較対象とするユーザの特性が似ているということである。

類似度は次式で定義する。

$$\text{類似度} = \frac{\text{整合した項目数}}{(\text{対象のユーザ}^5 \text{がチェックした項目数} + \text{比較対象とするユーザがチェックした項目数}) - \text{整合した項目数}}$$

ここで項目とは、ユーザの嗜好を細分化し特性ごとに分けたものである。会員登録の際にユーザは任意で自分の嗜好に当てはまる項目にチェックを入れる。また、項目が整合するとは二者<sup>5</sup>がチェックを入れた項目が同じであることを表す。

類似度の分子は二者の項目が整合した数とする。分母は二者がチェックした項目数の合計から、二者が整合した項目数を取り除くことによって、全部で何種類の項目を選択したかを表している。単に二者の項目がどの程度整合したかを比較するだけでは、全体の項目の種類に占める整合した項目数の割合が明確に表われてこない。例えば全体で20項目にチェックをつけたユーザAと全体で4項目にチェックをしたユーザBの整合数が4つだとすると、ユーザBは100%項目が整合したことになるが、ユーザAは20%項目が整合したというように二者の類似度に乖離が生じる。そのため、二者の項目がどの程度整合したかを比較するだけでは精度の高い類似度が求められない。しかし、この計算式にすることによって、二者が選択した項目の全体に占める整合した割合が正確に表れるため、類似度が求められる。

4.5.5 時間価値

時間価値とは、記事の時間に対する価値の変化を考慮するための値である。投稿された時間が新しい記事ほど評価が高く[14]、古い記事ほど評価が低いものとし、フィルタリングする。

<sup>5</sup>対象のユーザと比較対象とするユーザを指す。

中澤ゼミ

時間価値はデータベースから全体の経過時間を算出し、記事に応じて適した段階に区切りその範囲で点数化する。例えば、Sを4とすると4段階に区切られ、1-4の範囲で記事の取捨選択が行われる。

$$1 \text{ 段階当りの経過期間} = \frac{1 \text{ 番最後に投稿された記事の日付} - 1 \text{ 番最初に投稿された記事の日付}}{S}$$

4.6 フィルタリングのルール

フィルタリングのルールは、大きく分けて3つの評価値によって決定される。1つ目として、記事一つ一つに対する総合的な推薦度である加重平均推薦度である。これは、1-4の範囲で表されるために1点区切りで4段階の推薦度をコントロールバーで調節し記事を取捨選択できるものとする。例えば推薦度の目盛りが2の時は、加重平均推薦度が2の記事のみが提供されるというものである。例外として、評価をしたレビューア全体が同じ推薦度の場合、加重平均推薦度だけでなく信頼度同士を比較する。これによって同じ加重平均推薦度の記事であっても信頼度が高い記事から順に表示することができる。また、付加情報として記事の評価したレビューアの人数を表示する。

2つ目は、ユーザー一人一人の特性を他ユーザーと比較し特性がどの程度似ているかを表す類似度である。これは、0-1の範囲で表されるために0.2点区切りで5段階の類似度をコントロールバーで調節し記事を取捨選択できるものとする。例えば類似度の目盛りが5の時は、類似度が5の記事のみが提供される。

3つ目として、新しい記事ほど価値が高いことを表す時間価値が挙げられる。時間価値は記事が投稿された時間をデータベースに格納しているので、その時間の新しさや記事の特性によってS段階で価値をはかる。これは、1-Sの範囲で表される。これによって、S段階の時間価値をコントロールバーで調節し記事を取捨選択できるものとする。

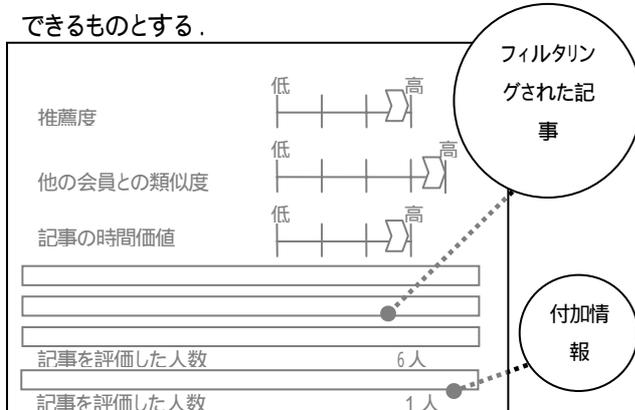


図 4: 記事出力画面

5 結び

本研究では、消費者がいかに信頼性の高い情報を得るかという点に重点を置き、情報の信頼性とユーザーの

特性に注目した口コミ掲示板システムを提案した。信頼性については記事とレビューアへの評価を加重平均推薦度として算出し、高い記事ほど信頼性が高いものとした。また、ユーザ特性を他のユーザと比較することにより、近い特性のユーザの記事を推薦するものとした。以上の点から、本システムでは、膨大な情報の中から信頼性の高い情報を取捨選択し利用者の望む情報を得る支援ができるものと考えられる。

消費者が商品選択のために、より多角的な情報を求めて生まれた口コミ掲示板は、現在では消費者間だけでなくメーカー側に消費者が自分で使った商品の評価に関する生の情報を伝える橋渡しの役割を果たしている。そのため、今後、口コミ掲示板によって企業は、よりマーケットインに近い思想で新しい戦略を考え、新たなビジネスチャンスをネットコミュニティの中から得る機会が増えてくると予想される。ネットコミュニティに参加する人々が増え企業に与える影響が大きくなることは、消費者の要望がサービスや商品に反映されるチャンスが増えることにつながると考える。そのため、消費者は、これまで以上に掲示板への書き込みなどに対して責任を持たなければならない。

参考文献

- [1]結城浩, perl でつくる CGI 入門・応用編・生きたホームページをつくるために, ソフトバンク社, 1998.
- [2]堀内明, Perl で学ぶゼロからのプログラミング, 日本評論社, 2002.
- [3]菅坂玉美, 横尾真, 寺野隆雄, 山口高平, "インターネットの知的情報技術・e ビジネスの理論と応用", 東京電機大学出版局, 2003.
- [4]http://www.kakaku.com/, 価格.com.
- [5]http://www.amazon.co.jp/, amazon.co.jp.
- [6]http://furima.rakuten.co.jp/, 楽天フリマ, 楽天.
- [7]http://www.auction.yahoo.co.jp/, Yahoo!オークション, Yahoo!Japan.
- [8]http://e-words.jp/, IT用語辞典 e-words.
- [9]http://www.sw.nec.co.jp/lecture/word/km/, ナレッジマネジメント, 5分でわかる IT, IT Square.
- [10]http://allabout.co.jp/career/cgiperl/, CGI・Perl, All About Japan.
- [11]http://perl.misty.ne.jp/, perl CGI 講座, ミスティーネット.
- [12]http://www.parkcity.ne.jp/~chaichan/src/perlmain.htm, CGI-Perl の基礎講座, Chaichan-World !.
- [13]http://www.mytools.net/cgigtips/index.html, CGI ってなあに? , MyTools.NET.
- [14]遠藤博人, 能登正人, "口コミモデルによる情報推薦システム", 電子情報通信学会誌, vol.103,pp13-18, 2003.
- [15] 神田 昌典, 口コミ伝染病 お客がお客を連れてくる実践プログラム, フォレスト出版,2001.
- [16]http://www.itmedia.co.jp/lifestyle/articles/0403/31/news072.html, "神の見えざる手"をサポート 価格.com の"クチコミ"パワー", lifestyle.
- [17]http://smallbiz.nikkeibp.co.jp/members/ITKEYWORD/20030423/103079/, 3分でわかる最新 IT キーワード, 日経 BP ネットワーク.