原著論文

Web ユーザレビューにおける評価情報の 時系列変化の可視化[†]

打田 裕樹^{*1}・吉川 大弘^{*1}・古橋 武^{*1}・平尾 英司^{*2}・井口 浩人^{*2}

近年、インターネットの普及を背景に、ブログやレビューなど、Web上で個人の日記や意見を気軽に書き込むことができる環境が整ってきている。本研究では、このように個人が発信するWeb上での自由記述情報を用いることで、製品などに対する評価解析を行うためのシステム構築を目指す。従来の評価解析システムでも評価表現の抽出や呈示がなされているが、個々の製品に対しての評価表現を把握することが目的であり、製品間の評価の比較や同製品での評価の時系列変化を把握する上では、不十分であると考えられる。本システムでは、MDSによりユーザレビュー間の関係性を把握した上で、文章中の語句どうしの関係を可視化するテキストマイニング手法HK Graphを用いて、ユーザが興味を持った対象商品に対するレビュー内容の視覚的な把握支援を行う。本論文では、Webユーザレビューから抽出された評価語句に基づいて、レビューにおける評価内容の時系列変化・類似性を視覚的に把握する手法を提案する。

キーワード：HK Graph, ユーザレビュー解析, テキストマイニング, 時系列変化, 評価語句

1. はじめに

近年、インターネットの普及を背景に、ブログやユーザレビュー、掲示板など、Web上で個人の日記や意見を気軽に書き込むことができる環境が整ってきている。それに伴い、自分の意見や考えをWeb上に公開するユーザや、興味があることにに対してWeb上で第三者からの意見を求めるユーザが増加してきている。しかし現在、Web上で起こっている情報爆発により、例えばある製品の評価に関する文章だけでも大量に存在する。そのためこれらすべてに目を通すことは困難である。そこで本研究では、個人が発するWeb上での自由記述情報を用いることで、製品などに対する評価解析を行うためのシステム構築を目指す。ユーザが製品に対して点数で評価値を与え、多数の投票の平均値をとることで、製品の定量的な評価を行うWebサイトなどはすでに存在するが、これらの情報だけでは、高い／低い評価の理由や、ユーザの真の声を解析することは困難である。自由記述形式であるユーザレビューを対象とすることで、定量的な評価の理由や、点数には表れないユーザの真の声、あるいは平均では埋もれてし

まうマイノリティの評価などの解析が可能となることが期待できる。しかし一方で、数値的には扱いにくいWeb上の膨大な量のユーザレビューの中から、有用な評価情報を効率的に取り出すことは容易ではない。

ユーザが家電製品、自動車、携帯電話などさまざまな製品への意見・感想を書き込むことができるWebサイトの1つとして、“価格.com”[1]が知られている。価格.comでは、製品ごとにユーザレビューが書き込まれており、何らかの製品に興味をもったユーザは、そのユーザレビューを読むことで、製品に対する具体的な評判情報を得ることが可能である。価格.comでは、それぞれの製品に対して、数十件から製品によっては数百件のユーザレビューが存在する。そのため、ユーザが製品の評判を知るためにこれらのユーザレビューをすべて読むとすると、多大な時間や労力を要することとなる。また、複数の製品を比較したい場合には、さらに多くのユーザレビューに目を通さなくてはならず、必要な評価情報を得ることが困難となる。また製品の中には、発売の時期などの影響等で、レビューが書き込まれた時期により評価の内容が変化しているものがあり、ユーザはこのような製品に対しては、変化する評価情報を時系列でとらえながらレビュー全体の傾向を把握する必要がある。

これまでも評価情報の抽出に関する研究は多くなされており、例えば文単位でどのような評価がされているのか、またそれが肯定的／否定的な評価であるかを判定することを目的とする手法[2][3]や語句単位で

[†] Visualization of Time Series Variation of Web User Review based on Evaluation Keywords

Yuki UCHIDA, Tomohiro YOSHIKAWA, Takeshi FURUHASHI, Eiji HIRAO and Hiroto IGUCHI

^{*1} 名古屋大学
Nagoya University

^{*2} 日本電気株式会社
NEC Corporation

評価を表すキーワードを特定し、それが肯定的／否定的な評価であるかを判定することを目的とする手法[4][5]などがある。対象とするWeb上のユーザーレビューは、形式的に書かれた文ではないため、複数の意味を持つ文も存在する可能性が考えられる。そこで本研究の目指すシステム(以下、本システム)では、後者のように語句単位で評価を表すキーワードを抽出することで、ユーザにより適切な評価情報の提供を行う。

また、情報の推移を時系列的に解析するシステムも多く提案されており、SHOOTI[7]やSTEND[8]などが挙げられる。SHOOTIでは、文書中のキーワードのburst度を算出することで時系列解析を行っている。しかしこのシステムでは、そのキーワードの出現度合しか見ていないため、詳細な評価の時系列変化の把握や複数のキーワードに関する評価内容の時系列変化の比較を行うことができない。またSTENDでは、文章集合からキーワードに関連する情報を収集し、そこから得られる動向をグラフとして提示しているが、ユーザーレビューにはそのような動向が含まれているとは考えにくい。本システムでは、キーワードに関する評価情報を用いて時系列解析を行うことで、そのキーワードに関する評価内容の時系列解析や複数のキーワードに関する評価内容の時系列解析を行う。前述の通り、製品の評判については、その発売からの時期や、新製品の発売のタイミングなどにより傾向が変わることがある。そこで本システムでは特に、月単位のユーザーレビューの類似性を可視化することで、時系列での製品評価の推移の視覚的な把握支援を行う。これにより、ある製品に対して、月単位でどのような項目が評価あるいは注目されているか、その推移を知ることができる。例えば携帯電話について、機種Aは全てのレビューをまとめて解析してしまうと、デザインに関する評価が多くされているように捉えてしまうが、時系列での推移を把握することで、実は発売当初に多く評価されていただけでその後はあまり評価されなくなったことがわかるなど、製品の評価内容の詳しい推移を把握することができる。また同時期に発売された機種Bを同時に解析することで、機種Bは機種Aとは逆に発売後しばらくしてからデザインに関して多く評価されるようになり、その後長く評価され続けていることがわかれば、デザインを基準に携帯電話を購入するユーザにとっては機種Bのほうが長く飽きずに使用できるといった情報も得ることができ、購入の意思決定に役立てることができる。またこれらの解析は、製品のメーカーにおけるマーケティング調査としても有効であると考えられる。

一方、文章データ中の語句同士の関係を、階層的なグラフ構造を用いて可視化するテキストマイニング手法としてHK Graph(Hierarchical Keyword Graph)[6]が提案されている。HK Graphでは、共起度情報に基づき、ユーザが入力した語句と関連性の高い語句をグラフ構造で可視化することで、膨大な文章データの概要を俯瞰的に把握することが可能となる。そこで本システムでは、ユーザーレビュー文章に対してHK Graphを適用することで、ユーザが興味のある製品に対する評判情報の視覚的な把握支援を行う。

本論文では、ユーザーレビュー解析支援システムの構築を目的として、Webユーザーレビューから抽出された“評価項目”[9][10]に基づき、レビューにおける時系列変化・類似性の可視化手法を提案する。具体的には、月ごとのレビューで用いられた評価項目によってレビュー間の類似度を定義し、多次元尺度構成法(Multi Dimensional Scaling : MDS)[11][12]を用いてそれらのレビュー間の関係を可視化する。これにより、製品を評価する上で注目された項目についての類似性を把握することが可能となる。MDSにより得られた可視空間上での時系列変化の形状を表す3つの指標に基づき、評判がどのように変化しているかを特徴づける。また、MDSによりユーザーレビュー間の関係を把握した上で、ユーザが興味を持った製品およびその時期についてのレビューを選択し、それらに含まれる評判情報をHK Graphを用いて呈示する。本論文では、前述した“価格.com”の携帯電話に関するレビューに提案手法を適用し、その有効性を示す。

2. HK Graph

2.1 HK Graphの特徴

文章から抽出した語句を、HK Graphにより可視化する。HK Graphとは、文章情報に付加された属性情報と、自由記述文章中の語句との共起度を利用して語句の抽出、階層的キーワードグラフの作成を行う手法である。HK Graphでは、ユーザが介在することにより、ユーザ自身が能動的に文章データの解析、探索を行う事ができる。

一方、文章中の語句の関係を可視化する手法として、KeyGraph[13][14]がある。KeyGraphとは、文章中の語句の出現頻度情報および共起情報を用いて語句を抽出し、グラフ構造で可視化することで、文章の重要な語句や事象の関係性を把握する手法である。KeyGraphでは、文章全体での語句の関係を呈示することは可能であるが、ユーザが主体的に語句を入力し、それらの語句と文章中の語句との関係を呈示するという適用方法は想定されていない。また他の可視化

手法としてACCENT[15]があり、ユーザによる語句の入力が可能となっている。このシステムでは、入力語句間の関係をグラフ構造で可視化し、語句間の関係を把握することができるが、さらに詳しい解析を行うためには再度グラフを構築し、グラフを見比べなければいけないため、全体傾向を把握するための俯瞰性に欠けている点があると考えられる。本研究では、ユーザが興味を持つ対象やキーワードに関する評判を知り、レビューの傾向を把握することを目的としているため、ユーザが能動的にグラフ構造や表示内容に関与することができるHK Graphを用いる。以下に、HK Graphのアルゴリズムを述べる。

2.2 語句への分割

初めに、文章データを語句に分割する。ここでは、対象となる文章データに対して、Cabocha[16]を用いて形態素解析および文節区切りを行い、語句への分割を行う。このとき、形態素に分割された語句に対して、その品詞情報を用いて、助詞、記号(句読点、括弧)、代名詞、接続詞、連体詞を解析に必要な語句として削除した上で、文節情報を用いて結合する。例えば「ボタン操作が簡単だ。」という文は、文節区切りおよび形態素解析を用いて分割すると以下のようになり、最終的に「ボタン操作」「簡単」という2つの語句に分割される。

- ・ 文節区切り
 {ボタン操作が、簡単だ。}
- ・ 形態素解析
 {ボタン、操作、が、簡単、だ、。}

2.3 ベースの選択

次に、ユーザが解析したいと考える内容、すなわちここではユーザが興味のある製品名を選択する。本論文では、属性情報として例えば携帯電話の機種名「N904i」などの製品名(複数選択可能)、文章データとしてそれらの製品について書かれたユーザレビューをそれぞれ用いる。ここで選択された語句は解析の基となるため、“ベース”と呼ぶ。2.4で述べる主ノードの抽出では、ベースとの共起度が計算され、表示する語句が抽出されるため、ベースが異なればキーワードグラフの表示結果も異なる。この点は、ユーザが能動的にグラフ構造に関与するというHK Graphの特徴の一つである。

2.4 主ノードの抽出

入力されたベースを利用して、文章データ中の語句から、ベースとの共起度が高い語句の抽出を行う。こ

こで、共起度の算出にはJaccard係数を用いる。ベースとして入力された語句を B_i 、抽出対象となる文章データ中のある語句を W_j 、とすると、 B_i と W_j の共起度は次式で表される。

$$Jaccard(B_i, W_j) = \frac{N_s(B_i \wedge W_j)}{N_s(B_i \vee W_j)} \quad (1)$$

ここで $N_s(X)$ は語句Xを含む文章の数である。ベースが2つ以上の場合、(1)式を用いて得られた各ベースとの共起度を掛け合わせた値を指標とし、ベース全体との共起度が高い語句(全結合語句)、特定の組み合わせのベースに対して共起度が高い語句(複結合語句)、各ベースに対して共起度が高い語句(単結合語句)の抽出をそれぞれ行う。ここで抽出された語句を“主ノード”と呼ぶ。

2.5 副ノードの抽出

“副ノード”には、ユーザが選択した任意の主ノードに対して関係性の高い語句が抽出される。つまり(1)式において、ベースを主ノードに置き換えたとき抽出される語句となる。ただしシステムの画面上ではベースと主ノードのみが表示され、ユーザが、主ノードを選択(クリック)したときにその主ノードに対する副ノードが表示される。また、それぞれの副ノードについて、対象となる主ノードの他に、最も共起度が高い他の主ノードとのリンクを表示する。

2.6 階層的キーワードグラフの呈示

図1に、HK Graphにより作成された階層的キーワードグラフのイメージ図を示す。図において、 $B_1 \sim B_3$ がベース、 A_1 が全結合語句、 $M_{12} \sim M_{23}$ が複結合語句、 $S_{11} \sim S_{32}$ が単結合語句、 Sub_1, Sub_2 が S_{32} に対して共起度の高い副ノードである。このように、ベースと主ノードはリンクの接続により関係が表され、語句同士の共起度の値はリンクの太さにより表されている。

図1 HK Graphのイメージ

3. 提案手法

ここでは、ユーザレビュー文章中から対象の評価の際注目した点を表す評価項目を抽出する方法、また抽出された語句に基づき、レビュー同士の関係を視覚的に把握する方法を説明する。

3.1 評価項目の抽出

初めに、日本語係り受け解析ツール“Cabocha”[13]を適用し、得られた係り受け情報を用いて“評価項目”を抽出する。図2に、携帯電話に対する評価文章の例を示す。ここで評価項目とは、「デザイン」、「音質」など、ユーザが製品などを評価する際に着目した点である。本論文では、形態素情報に基づき、図2の「かっこいい」など、形容詞／形容動詞に係っている名詞・未知語を評価項目として抽出する。実際には、評価を表す語としては、形容詞／形容動詞だけでなく、名詞や動詞に係る語も存在するが、逆に名詞や動詞に係る語の多くは評価項目としては適切ではないため、ここでは形容詞／形容動詞に限定した。また、未知語を対象としたのは、「アプリ」や「ワンセグ」などの語句が評価項目となりうるためである。なお、この方法による評価項目の抽出性能などの検討は[9][10]で示す。现阶段では、抽出性能が十分であるとはいえないが、抽出方法の改良については今後の課題とし、本論文では上述の方法で抽出された評価項目を用いる。

3.2 評価項目を用いたHKGraph

図3に、3.1で抽出した評価項目を、2.4で示した主ノードとしてHKGraphにより提示した例を示す。図では、携帯電話の3つの製品をベースとして提示している。この図から、3つの製品に共通して「ワンセグ」や「レスポンス」、「バッテリー」について評価されている(全結合語句)ことや、「ヨコモーション」、「ニューロポインタ」など、各製品特有の機能などがそれぞれの製品で評価されている(単結合語句)こと、ある製品は特に「画質」や「音」に注目して評価されていることなどがわかる。

このようにHK Graphを用いることで、製品に対する評判の把握や、製品間の評判を比較することは可能であるが、それらの時間的変化を把握することは困難である。そこで本論文では、各製品のレビューを月単位に分けてHK Graphを適用する。これにより、1つの製品に対する評判の時系列変化や、複数の製品における評判の時間的な推移を比較することが可能となると考えられる。なお本論文では、携帯電話のモデルチェンジのタイミングが半年～1年程度であること、

図2 携帯電話に対する評価情報を含む文章の例

図3 評価項目を主ノードとしたHK Graphの呈示例

また平均レビュー数を考慮して時間的な区切りを月単位としたが、この区分の単位については、対象とする製品の製品寿命や解析の目的によって適切な区分単位を用いる必要がある。一定の区分で分けるのではなく、新製品の発売直後は短く、その後徐々に長くしていく区切り方も有効であると考えられる。一方、レビューを月単位で分割することで、ベースの候補となる製品や時期の組み合わせが膨大となる。そこで、各製品・月のレビューの類似関係を可視化し、各製品に対する評価の時間的変化や製品ごとの評判の類似性などの全体像を把握した上で、ユーザが興味を持ち、HK Graphで具体的な解析を行う製品や時期を抽出する。また、この全体像の把握とHK Graphによる具体的な解析との連携により、ユーザとシステム間のインタラクション支援[17]を行う。以下に、レビュー間の類似度の定義方法、および指標を用いたレビュー間の時系列特徴の定量化方法を示す。

3.3 レビュー間の類似度の定義

HK Graphを用いたレビュー解析を行うにあたり、各製品について、レビューで評価されている評価項目(評価着目点)の時間的変化や、製品ごとの評価項目の類似性を、大まかに把握することを目的とし、各月レビューの評価項目の類似性を可視化する。ここでははじめに、2つのレビューA、B間の類似度を定義する。

各レビューに対し、3.1で抽出された評価項目のレビュー内での出現頻度を算出する。レビューAまたはBから抽出された評価項目を X_i 、各レビュー内での X_i の出現頻度を x_i とし、レビューAとBの対におけるAの評価項目の出現頻度行列を以下のように表す。

$$X^{AB} = (x_i | X_i \in A \cup B) \quad (2)$$

この出現頻度行列は、 $n_{AB} \times 1$ 行列となり、行要素はレビューAまたはBから抽出された各評価項目に対するレビューA内での出現頻度であり、 n_{AB} はレビューAまたはBから抽出された評価項目数を表す。(2)式のようにレビューA、Bに対する出現頻度行列を求め、各行列 X^{AB} 、 X^{BA} を用いてレビューAとレビューBとの類似度、 $Sim(A, B)$ を以下のように定義する。

$$Sim(A, B) = \frac{X^{AB} \cdot X^{BA}}{|X^{AB}| |X^{BA}|} = \frac{\sum a_i b_i}{\sqrt{\sum a_i^2} \sqrt{\sum b_i^2}} \quad (3)$$

このとき、 a_i はAの要素を表し、 b_i はBの要素を表す。(3)式により定義された類似度を全レビュー間で求めることで、レビュー同士の関係性を数値的に表すことが可能となる。本論文では、月単位でのレビューに対し、上述の方法で類似度を求めることで、ある製品における時間的な評価の推移を把握する。

3.4 レビュー間の関係性の可視化

3.3で得られたレビュー間の類似度に対し、MDSを用いて全レビュー間の関係性を可視化する。MDSにより可視化された2次元マップより、レビュー同士の評価項目の類似関係を把握することができる。特に月単位でのレビューに対して適用することで、評価の時系列変化を把握することが可能となる。図4に、製品Aのある月から3ヶ月分のレビュー(A_1 , A_2 , A_3)に対して(3)式により類似度を求め、MDSを用いて可視

化したときのイメージ図を示す。マップ上で、各点の距離が大きいほど類似度が低く、距離が小さいほど類似度が高いことを示している。図において、 A_c は $A_1 \sim A_3$ の2次元平面上での重心、 A_a は3ヶ月のレビューをすべてまとめて1つのレビューとしたとき(“総括レビュー”と呼ぶ)、MDSによりプロットされた点を示す。本論文では、これらの点の分布形状を表す以下に示す3つの指標により、この製品に対するユーザーレビューの時系列変化を特徴づける。

指標1：総括レビューの偏り

A_c と A_a との距離 d_{ca} が大きいほど、総括レビューの内容がある月のレビューに偏っていることを表している。例えば図4のような例では、製品Aに対する(すべての月の)レビューを解析することは、実は A_3 の月のレビューのみを見ていることに近いことを意味しており、本来異なる評価がされていた A_1 や A_2 の内容を見落としてしまう可能性がある。そこで、総括レビューの偏りを以下のように定義する。

$$(\text{総括レビューの偏り}) = d_{ca} - d_{ia} \quad (4)$$

ただし d_{ia} は、総括レビュー A_a と、 A_a と最も近い各月レビュー A_i との距離を表す。

指標2：各月レビューの広がり

(5)式に示す、各月の重心からの距離(図4では $d_1 \sim d_3$)の平均は、各月のレビューの類似度を表している。それぞれの月が大きく広がってプロットされているほど、月ごとに評価されている内容が異なり、また逆に密集しているほど、時間がたっても評価の内容は類似していることを示している。

$$(\text{各月レビューの広がり}) = \frac{1}{N_R} \sum_i d_i \quad (5)$$

ただし d_i は重心 A_c からの各月レビューへの距離、 N_R は各月レビューの総数を表す(図4では $N_R = 3$)。

指標3：各月レビューの偏り

各月のレビューの偏りとして、例えば他の月に対してある月が1点のみ離れてプロットされている場合、その月のみに於いて、他の月にはない特別な評価がされていると考えられる。そこで、各月レビューの偏りを以下のように定義する。

$$(\text{各月レビューの偏り}) = \max_i \left\{ \frac{1}{N_R - 1} \sum_j (d_{ij} - \bar{d})(i \neq j) \right\} \quad (6)$$

図4 MDSによる可視化結果のイメージ

$$\bar{d} = \frac{1}{N_R C_2} \sum_{i=2}^{N_R} \sum_{j=1}^{i-1} d_{ij} \quad (7)$$

ただし d_{ij} は、ある月 i と他の月 j のレビューとの距離を表す。すなわち式(6)は、他の月から最も離れた月のレビューとその他の月のレビューとの距離の偏差を表す。なお $i \neq j$ より、レビュー i に対するレビューの数は $N_R - 1$ となる。

3.5 ユーザレビュー解析支援システム

3.1から3.4で述べた方法を、システムとして実装した。図5にシステムのインターフェイスを示す。このシステムは、画面左に評価情報を呈示したHK Graph、右にMDSによるマッピング結果の2つの部分からなる。

はじめに、ユーザはMDSによるマッピング結果をもとにレビュー間の類似性を把握する。MDSによるマッピング結果では、プロットされた点の距離関係によりレビュー間の類似性を把握することが可能である。また、図5の右下にあるウィンドウにおいて着目する製品を選択し、時系列情報表示ボタンをクリックすることで、MDSによるマッピング結果上でその製品の総括レビュー、また月ごとのレビューおよび重心点が色付けされ、さらに時系列変化の経緯が線で結ばれて表示されるとともに、3.4で示した3つの指標の値が表示される。これによりユーザは、その製品に関する総括レビューの偏りや各月レビューの広がりおよび偏りを把握しながら、解析するベースを決定することができる。

次に、ユーザはHK Graphにより、決定したベースに関するレビューを解析する。先の図8の例に示すように、HK Graphでは、レビューから抽出した評価情報を階層的に呈示するため、評価項目と評価語[10]の関係を保ちつつユーザに呈示することができる。図8

では、主ノードが評価項目であり、副ノードが評価語となっている。よって、ユーザが興味のある評価項目についての評価内容を階層構造を利用して知ることができるため、ユーザの意思を反映した、インタラクティブな解析を行うことが可能となる。また、HK Graphでは複数のベースに対して同時に解析ができるため、呈示された評価項目によりレビューの概要や時系列的な傾向を視覚的に把握することができる。

4. 実験

本論文では、Webサイト“価格.com”の携帯電話に関するユーザレビュー(2007年5月～2008年5月)に提案手法を適用した。本実験では、MDSによる可視化結果とHK Graphとの対応関係に関する実験とユーザの情報獲得や意思決定の支援に対する本システムの有効性を示す実験を行った。前者については、4.1、4.2で示し、また後者については、4.3で示す。

4.1 レビュー間の類似度に基づくMDSによる可視化

本実験では、“価格.com”の携帯電話に関するユーザレビューの中で、ドコモ10機種に関するレビュー750件を対象とした。初めに、3.1で述べた方法により評価項目の抽出を行ったところ、6878語の名詞・未知語から920語の評価項目が抽出された。この6878語の名詞・未知語から評価項目として適切な語句を入手により判別したところ、823語の適切な語句が含まれていた。さらに提案手法により抽出された評価項目内には、390語の適切な語句が含まれていたことから、提案手法による評価項目の抽出性能は、正答率が42.4%、網羅率が47.4%であった。また、全ての名詞・未知語を用いた場合の抽出性能は、正答率が12.0%、網羅率が100%である。相互情報量[18]を用いて比較すると、全名詞・未知語抽出の場合の相互情報量は0、提案手法では0.071となった。920語の抽出された評価項目を用いて、(3)式により全レビュー間の類似度を算出した。図6に、得られた類似度に基づき、MDSにより可視化を行った結果を示す。図において、機種名の後に付いている数字は、対象となるレビューの月を表し、数字が付いていない点はその機種に関するすべての月の総括レビューを表している。

HK Graphでは、各ベースあたりの呈示語句数をユーザが選択できる。ただし、このとき呈示される語句(主ノード=評価項目)は、各ベースに関するレビューにおいて、3.1で述べた形容詞/形容動詞への係り頻度が高い語句から順に選択され、他のベースで抽出された主ノード(評価項目)との共通性により、2.4で述べた全結合語句・複結合語句・単結合語句が決定され

図5 ユーザレビュー解析支援システムのイメージ

図6 類似度を用いたMDSによる可視化結果

る。これにより、全結合語句は、すべての月にわたって評価に用いられた語句であることを意味し、単結合語句はそのベースとなる月でのみ主に用いられた評価項目を表している。ベースあたりの呈示語句数は同じであるため、例えば単結合語句の割合が高いことは、その月が他の月とは異なる評価がなされたことを示している。

4.1.1 指標1：総括レビューの偏り

初めに、3.4で示した指標1に対する妥当性を検証する。総括レビューの内容と類似点が多くなると、総括レビューと結合する語句の割合が高くなると考えられる。そこで、HK Graph上での総括レビューとの共通語句の割合を以下のように定義し、指標1((4)式)との相関により、指標1に対する検証を行う。

総括との共通語句割合

$$= \frac{\text{総括レビューとの結合語句数}}{\text{呈示語句数}} \quad (8)$$

MDSによる結果には初期値依存性があるため、10回試行の平均により(4)式の値を求めた。また共通語句の割合を計算する際には、呈示語句数による影響を小さくするため、ベースあたりの呈示語句数を5～30語まで5語ずつ増やし、そのときの平均を用いた。ただし月によっては、レビューが少なく、抽出された評価項目が30語に満たない場合もあったため、その場合は呈示語句数を5語～抽出された評価項目数までとした。

図7に、本実験で用いた10機種について、(4)式の値に対する総括との共通語句割合を示す。図7におい

て、指標1と総括との共通語句割合との相関係数は0.696、P値は0.0252であった。これにより、指標1が総括レビューの偏りを表しており、この値が大きいことは、その機種の評判がある月に偏っていることを示していると考えられる。

また図6において“F904i”に着目すると、総括レビュー“F904i”の点が重心から離れ、“F904i-6”の近くにプロットされていることがわかる。図8に、“F904i”に関する6～8月の月ごとのレビューと、それらを総括したレビューをそれぞれベースとし、ベースあたりの呈示語句数を20語としたときのHK Graphによる出力結果を示す。図8から、総括レビューと6月のレビューとの複結合語句が多く、6月の単結合語句がほとんど呈示されていないことがわかる。すなわち、F904iに関するレビューとして解析を行った場合、6月に書かれたレビューに偏って評判を把握することとなり、実際には7月に書き込まれた「フルブラウザ使い易い」や「バッテリー餅みたい、減らない」といった評価や、8月の「変換-良く」、「ヨコモーション-優秀」といった評価、そして評価項目のそれ

図7 指標1に対する総括との共通語句割合の変化

図8 “F904i”に関するHK Graphの呈示結果

ら時間的変化を見落としてしまう可能性がある．図8の結果から，総括の偏りを把握した上でHK Graphで解析を行うことで，例えばF904iにおいてはマイナーな評価項目を中心に評判を解析することなどが可能となる．

4.1.2 指標2：各月レビューの広がり

4.1.1で述べた通り，月ごとのレビュー内容が異なるほど，それぞれの月の単結合語句の割合が高くなると考えられる．そこでここでは，HK Graph上での単結合語句の割合を以下のように定義し，指標2（(5)式）との相関により，指標2に対する検証を行う．

$$\text{単結合語句割合} = \frac{\text{全ベースの単結合語句数}}{\text{ベース数} \times \text{呈示語句数}} \quad (9)$$

4.1.1と同様，(5)式の値は10回試行の平均により求め，また単結合語句の割合はベースあたりの呈示語句数を5～30語まで5語ずつ増やし，そのときの平均を用いた．

図9に，本実験で用いた10機種について，(5)式の値に対する単結合語句割合を示す．図9において，指標2と単結合語句割合との相関係数は0.746であった．このとき，P値が0.0133であったため，指標2と単結合語句割合には相関関係があったといえる．このことは，MDSでプロットされる月別レビューの広がり，月ごとの評価項目の多様性に対応していることを示しており，指標2の妥当性が示されたと考えられる．各月のレビューの広がりにより，各対象（製品）に対する評価の時間的な類似性が把握できることや，それらを製品ごとに比較できることで，評価が安定している機種や，評価対象として着目される機能などが大きく変わっていく機種などを解析することができ，レビューによる評価解析を行う上で有効であると考えられる．

図10(a)(b)に，それぞれN904i, D904iに関する6～8月のレビューに対し，HK Graphを適用した結果を示す．指標2の値は，(a)N904iが0.208，(b)D904iが0.452である．上述の通り，指標2の値の差により，全結合語句や単結合語句の割合が異なっていることがわかる．また，N904iでは，「ボタン」，「画面」，「音楽」など，比較的基本的な機能について3ヶ月通して評価がされているのに対し，D904iでは，6月には「着信音」や「スピーカー」，7月には「バッテリー」や「音楽」，8月には「メール」や「メニュー」など，着目される機能が推移していることがわかる．

図9 指標2に対する単結合語句割合の変化

(a) N904i(指標2：0.208)

(b) D904i(指標2：0.452)

図10 HK Graphの呈示結果

4.1.3 指標3：各月レビューの偏り

ここでは、3.4の式(6)で対象となった月(ここでは最偏月と呼ぶ)のレビューについて、HK Graph上での単結合語句の割合を以下のように定義し、指標3((6)式)との相関により、指標3に対する検証を行う。

$$\text{最偏月の単結合語句割合} = \frac{\text{単結合語句数}}{\text{呈示語句数}} \quad (10)$$

ここでも4.1.1と同様、(6)式の値は10回試行の平均により求め、また単結合語句の割合はベースあたりの呈示語句数を5～30語まで5語ずつ増やし、そのときの平均を用いた。

図11に、本実験で用いた10機種について、(6)式の値に対する最偏月の単結合割合を示す。図11において、指標3と最偏月の単結合語句割合との相関係数は0.662、p値は0.0372であった。これにより、指標3の値が大きいことは、他の月にはない偏った評価がされた月が存在していることを表していると考えられる。

また図6に見られるように、D905iの4月のレビューが他の月から大きく離れてプロットされている。そこで図12に、D905iの2月～4月のレビューに対し、

図11 指標3に対する最偏月の単結合語句割合の変化

図12 “D905i”(2～4月)に関するHK Graphの呈示結果

HK Graphを適用した結果を示す。HK Graphの呈示結果についても、極端ではないが、“D904i-4”に対する単結合語句が2、3月のものに比べて多いことがわかる。特に2、3月の単結合語句には、評価項目としては適切ではないと考えられる「au」「評価」「撤退」「不満」や、同じ意味の語句が抽出されている「スピーカ」「スピードセクター」などが含まれているため、実質的には4月のレビューについて評価項目が大きく偏っているといえる。このことから、D905iは4月になってそれまでにあまりされていない「起動-早くなつ、遅い」、「イルミネーション-ある」、「カメラ-よく」などの評価がされるようになったことがわかり、D905iが発売されてからの期間や、他機種の発売など、評価が変化した背景を考察することで、より深い評判解析が可能となると考えられる。

4.2 複数の製品評価の時系列推移

4.1では、各製品に対しての時系列での評価の推移を検討した。ここでは、複数の製品に対する評判を、評価された時期を同期して解析する。本実験で用いた携帯電話の機種のうち、F905i、N905iなど905iシリーズはほぼ同時期に発売されているため、それらの製品の同月のレビューを解析することで、製品同士の比較を行う。例えば図6において905iシリーズの4月のレビューに着目すると、総括レビュー“4月総括”の点は重心からはあまり離れておらず、905iシリーズの5つの製品が互いに離れてプロットされていることがわかる。このように複数の製品の同月のレビューの点に着目することで、その月におけるレビュー同士の評価項目の類似関係を把握することができる。図13に、905iシリーズ5機種に関する、(a)発売当初、(b)発売一ヶ月後、(c)発売五ヶ月後のレビューに対してHKGraphを適用した結果を示す。図13(a)では、それぞれメーカーがアピールする機能に対する評価(「タッチクルーザー使える」、「Bluetooth-搭載し」、「スピードセクター-いい、使い心地」など)が機種ごとにされているが、一ヶ月後のレビューに対する結果である図13(b)では、機種ごとの差が極端に少なくなり、一般的な携帯電話の機能に対する評価(「デザイナー-好き、良い、気に入る、カッコイイ、上品」、「レスポンス-よい、最高、モッサリする、サクサク」など)が多くなっていることが把握できる。このことから、ユーザは、発売当初にはメーカーがアピールする機能に対して興味を持つが、しばらくすると携帯電話の一般的な機能に着目するようになることが推測される。また図13(c)では、発売当初には着目されていなかった、メーカー固有の機能に対する評価(「PCライクキー

(a) 発売当初

(b) 発売一ヶ月後

(c) 発売五ヶ月後

図13 905iシリーズに対するHK Graphの呈示結果

使ってる」)されるようになっていことがわかる。さらに、図13(b)で全結合語句であった項目が図13(c)では複結合語句(「カメラよく、きれい」「操作一慣れ、満足し」)になっていることから、これらの項目に対して特徴を持つ製品に評価が推移していったことも把握できる。このように、複数の製品に対して時系列での評価の違いを見ることで、これから製品を買うユーザにとっては、それぞれの製品の特徴を捉えながらレビューを見ることが可能となる。またメーカーにとっては、自社のアピールしている機能がユーザにどの程度着目されているか、他の製品に対してどのような機能において特徴差があるのかといったことを把握することができ、マーケティング戦略などに活用することができると思われる。

4.3 被験者による評価実験

本節では、ユーザの情報獲得や意思決定の支援に対する本システムの有効性を検証するために、被験者による評価実験を行う。ここでは、[7][8]のように、文章や語句の出現頻度の推移に基づいて時系列変化を把握しながら、評価情報を把握するシステムを従来手法として、提案手法との比較を行った。具体的には、すべてのレビュー文章を一覧表示したものに加え、各製品の月ごとのレビュー数の変化を示したグラフ、および各月のレビュー内での出現頻度の高いキーワード上位10個を呈示した。

本実験では、被験者として大学生12名(20代・男性)の協力を得た。被験者には、極力客観的に評価をするよう指示した上で、提案手法と従来手法それぞれを用いて、これから購入する製品の検討、決定を行うことを想定してレビューを見てもらった。6名には先に提案手法を、残りの6名には先に従来手法を用いて実験を行った。本実験では、4.1で示した10機種の携帯電話のうち、簡易化のため905iシリーズの5機種に関するレビューの解析を行ってもらった。定量的評価として、購入製品決定までに要した時間の計測を行うとともに、定性的評価のために実験後に以下の3つ質問に答えてもらった。なお、情報を獲得する量が増え、その結果意思決定までに要する時間が増えることもあるため、購入製品決定までに要した時間の比較は参考程度の情報である。

- 質問1 レビュー内容の把握・比較の行いやすさ
- 質問2 時系列情報の有用性
- 質問3 意思決定支援としての有効性

これらの質問に対して、5段階評価により、従来手法

が優れていた(1点)から提案手法が優れていた(5点)として回答してもらった。

実験の結果、従来手法を用いた場合における購入製品決定までの平均時間は27分53秒(±8分39秒)、提案手法を用いた場合の平均時間は21分26秒(±5分55秒)であった。()内はそれぞれ標準偏差である。平均値の差に対するt検定を行ったところ、p値が0.11となり、購入製品決定までの平均時間に有意な差は見られなかった。また上述の各質問に関する平均評点は、質問1が4.3(±1.1)、質問2が3.6(±0.8)、質問3が4.0(±0.7)であった。本評価実験結果からのみでは、提案手法に対する客観的な評価を行うことは困難ではあるものの、提案手法を用いてレビュー内容の把握や製品比較を行うことで、ユーザの意思決定を支援できる可能性が示唆されたと考えられる。

また、本実験において被験者から、「右の図(MDSによる可視化)の意味が最初はわからなかったが、使っていくうちに少しずつ理解できた。」や、「HK Graphの表示を見てから指標の値を見るときは思った。」といった意見が聞かれたため、同じ被験者(少し慣れた状態)に再度提案システムを利用してもらい、特に3.4で示した各指標についての評価を行ってもらった。以下の質問について、それぞれ5段階評価(質問①～③:役に立たない(1点)～役に立った(5点)、質問④:使いにくくなった(1点)～使いやすくなった(5点))により回答してもらった。

質問① 指標1の使いやすさ

質問② 指標2の使いやすさ

質問③ 指標3の使いやすさ

質問④ 指標4前回の実験に比べ、システムが使いやすくなったか

その結果、各質問について、質問①4.9(±0.3)、質問②4.0(±0.9)、質問③3.6(±0.8)、質問④4.0(±0.4)という評価を得た。ユーザにとって本指標がおおむね有効であったと考えられる。ただし、最初の実験時においては、指標についての評価はもっと低かったと考えられ、これらの指標を有効に用いるためにはある程度の“慣れ”が必要となると考えられる。

追加実験において、「慣れてくると使いやすい。」、「指標を参考にすることで、着眼点がわかりやすくなった。」、「指標1と3の違いがあまりわからない。」といった意見が聞かれた。質問の評価が低かったこととあわせ、特に指標3について検討していく必要があると考えられる。

5. まとめ

本論文では、Webユーザレビューから抽出された評価項目に基づき、レビューの時系列変化・類似性の可視化手法を提案した。月ごとのレビューで用いられた評価項目によってレビュー間の類似度を定義し、MDSを用いてそれらのレビュー間の関係を可視化した。提案手法により、製品を評価する上で注目された項目についての類似性を把握することが可能となることを示した。また本論文では、MDSにより得られた可視空間上での時系列変化の形状を表す3つの指標に基づき、評判がどのように変化しているかの特徴づけを行った。“価格.com”の携帯電話に関するレビューに提案手法を適用し、MDSによる可視化結果とHK Graphによる呈示結果との対応関係を確認した。さらに、複数の製品に対する評判を、評価された時期を同期して解析することで、ユーザやメーカーにとって有益な製品の特徴把握が行えることを示した。

参考文献

- [1] 価格.com: <http://bbs.kakaku.com/bbs/>
- [2] Peter D. Turney. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings 40th Annual Meeting of the Association for Computational Linguistics (ACL'02), pp.417-424, 2002
- [3] 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”, 第18回人工知能学会全国大会, 2004
- [4] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: “テキストマイニングによる評価表現の収集”, 情報処理学会研究報告, 2003-NL-154, pp.77-84, 2003
- [5] 鈴木泰裕, 高村大也, 奥村学: “Weblogを対象とした評価表現抽出”, 人工知能学会セマンティックウェブとオントロジー研究会, SIG-SWO-A401-02, 2004
- [6] 岡部貴博, 吉川大弘, 古橋武: “インシデントレポート解析のための多重接続型階層的テキストマイニング手法の提案”, 第22回ファジィシステムシンポジウム, pp.211-214, 2006
- [7] 藤木稔明, 南野朋之, 鈴木泰裕, 奥村学: “document streamにおけるburstの発見”, 情報処理学会研究報告, 2003-NL-160, pp.85-92, 2004
- [8] 加藤恒昭, 松下光範: “情報編集 (Information Compilation) の基盤技術”, 2006年度人工知能学会全国大会 (第20回) 論文集, 1D3-02, 2006
- [9] Yuki Uchida, Tomohiro Yoshikawa, Takeshi Furuhashi, Eiji Hirao, Hiroto Iguchi: “Evaluation of Products by Analysis of User-Review using HK Graph”, Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced Intelligent Systems, pp.376-379, 2008
- [10] Tomohiro Yoshikawa, Yuki Uchida, Takeshi Furuhashi, Eiji Hirao, Hiroto Iguchi: “Extraction of

- Evaluation Keywords for Analyzing Product Evaluation in User-Reviews Using Hierarchical Keyword Graph", Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.13, No.4, pp.457-462, 2009
- [11] Norimasa Hayashida, Hideyuki Takagi: "Visualized IEC: Interactive Evolutionary Computation with Multidimensional Data Visualization", IEEE International Conference on Industrial Electronics, Control and Instrumentation (IECON2000), Vol.4, pp.2738-2743, 2000
- [12] Takeshi Yamada, Kazumi Saito, Naonori Ueda, "Cross-Entropy Directed Embedding of Network Data", Proceedings of the Twentieth International Conference on Machine Learning, pp.832-839, 2003
- [13] 大澤幸生, ネルス E. ベンソン, 谷内田正彦: "語の共起グラフの分割・統合によるキーワード抽出", 電子情報通信学会誌, Vol.J82-D-1, No. 2, pp.391-400, 1999
- [14] 久代紀之, 大澤幸生: "多次元ヒヤリングと階層的な要求統合プロセスによる要求獲得手法", 情報処理学会論文誌, Vol.47, No.10, pp.2909-2916, 2006
- [15] 三末和男, 渡部勇: "テキストマイニングのための連想関係の可視化技術", 情報処理学会第55回情報学基礎研究会資料, Vol.99, pp.65-72, 1999
- [16] 工藤拓, 松本裕治: "チャンキングの段階適用による係り受け解析", 情報処理学会論文誌, Vol.43, No. 6, pp.1834-1842, 2002
- [17] 稲村博央, 庄司裕子: "XBRLを利用した財務分析システムとそのインタラクション支援", 日本知能情報ファジィ学会誌, Vol.21, No. 3, pp.338-347, 2009
- [18] C.E.Shannon: "A Mathematical Theory of Communication", The Bell System Technical Journal, Vol.27, pp.379-423, 623-656, 1948
- (2009年7月29日 受付)
(2010年3月17日 採録)
- [問い合わせ先]
〒464-8603 名古屋市中千種区不老町
名古屋大学 大学院工学研究科 計算理工学専攻
吉川 大弘
TEL: 052-789-3167
FAX: 052-789-3166
E-mail: yoshikawa@cse.nagoya-u.ac.jp

著者紹介

うちだ ゆうき
打田 祐樹 [学生会員]

2008年名古屋大学工学部電気電子情報工学科卒業。2010年名古屋大学大学院工学研究科博士前期課程計算理工学専攻修了。同年株式会社デンソー入社。現在に至る。在学中は主としてテキストマイニングに関する研究に従事。

よしかわ ともひろ
吉川 大弘 [正会員]

1997年名古屋大学大学院博士課程修了。同年カリフォルニア大学バークレー校ソフトコンピューティング研究所客員研究員。1998年三重大学工学部助手。2005年名古屋大学大学院工学研究科COE特任准教授。2006年10月同研究科准教授。現在に至る。主としてソフトコンピューティングとその応用に関する研究に従事。博士(工学)

ふるはし たけし
古橋 武 [正会員]

昭和60年名古屋大学大学院工学研究科博士後期課程電気系専攻修了。工学博士。平成16年名古屋大学大学院工学研究科計算理工学専攻教授。現在に至る。ソフトコンピューティング、感性工学に関する研究に従事。平成8年日本ファジィ学会論文賞受賞。IEEE, 日本知能情報ファジィ学会, 電気学会等の会員

ひらお えいじ
平尾 英司 [非会員]

1999年横浜国立大学大学院工学研究科博士前期課程物質工学専攻修了。同年日本電気株式会社入社。現在に至る。主として、アンケート設計・分析技術の開発、ITシステムの省エネ技術の開発などに関する研究に従事。

いぐち ひろと
井口 浩人 [非会員]

1982年名古屋大学大学院工学研究科修士課程電気系専攻修了。同年日本電気株式会社入社。現在に至る。光伝送装置の開発、環境・エネルギー研究、サービス研究に従事。電子情報通信学会会員。

Visualization of Time Series Variation of Web User Review based on Evaluation Keywords

by

**Yuki UCHIDA, Tomohiro YOSHIKAWA, Takeshi FURUHASHI,
Eiji HIRAO and Hiroto IGUCHI**

Abstract :

Recently, the number of the web sites such as Web user reviews and web logs where users can express their private ideas and opinions is increasing. However, it is very difficult to read whole the reviews on the internet. The conventional systems for analysis of reviews can extract evaluation information, but they are not enough to analyze time-series variation of them. This study tries to develop a review analysis system which shows the relationship among user reviews by MDS and shows evaluation information for products using HK Graph which can visualize the relationship between words with hierarchical network structure based on the co-occurrence information for the keyword graph. This paper proposes the visualization method to analyze time-series variation of Web user reviews based on the similarity of extracted evaluation keywords.

Keywords : HK Graph, User Review Analysis, Text Mining, Time-Series Analysis, Evaluation Keywords

Contact Address : **Tomohiro YOSHIKAWA**

Dept. of Computational Science and Engineering, Graduate School of Engineering, Nagoya University

Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN

TEL : 052-789-3167

FAX : 052-789-3166

E-mail : yoshikawa@cse.nagoya-u.ac.jp