

単語特性辞書を用いた電子掲示板内のコミュニケーションにおける 雰囲気の良い悪いおよび変化の同定

Mood Identification in Bulletin Board System with Word Characteristic Dictionary

濱岡 秀平

Shuheji Hamaoka

砂山 渡

Wataru Sunayama

広島市立大学大学院情報科学研究科

Graduate School of Information Sciences, Hiroshima City University

Recently, bulletin board systems are widely used as communication tools in the Internet. However, it takes too much time to find flourish communication in a good mood when we try to find meaningful information. In this paper, a system that identifies moods of communication held in a bulletin board system is proposed. The system use two word sets, one is for a good mood and the other is for a bad mood, to identify mood of each communication. By using this system, users can see which communication is in a good mood or bad mood, and can select a good one.

1. はじめに

近年、インターネットの普及により、Web 上で膨大な情報を Web 上から取得できるようになった。Web 上から情報を取得する手段として、不特定多数のユーザと情報のやりとりが出来る電子掲示板が広く利用されている。多くの電子掲示板では、匿名性のもが多く誰でも参加可能であり互いに面識のない利用者同士が自由に意見を書き込み情報交換や会話・議論を行う事ができ、多数の様々な意見のやりとりが盛んに行われている。そのため、電子掲示板内でのやりとりの中には、ユーザが互いに共感しあったりする雰囲気の良い部分と互いに揶揄しあったり誹謗中傷が発生したりする雰囲気の悪い部分が存在する。その雰囲気を読み取るには、実際にユーザが閲覧する必要がある。しかし、ユーザは雰囲気の良い部分の書き込みを中心に閲覧したり、雰囲気の悪い部分の書き込みを避けて閲覧したい場合がある。

そこで本研究では、電子掲示板内において雰囲気の良い部分と悪い部分を同定する事を目的とする。さらに、スレッド(一つの話題に関するユーザの書き込みの集合)内の雰囲気の流れをグラフで視覚化し、またスレッド全体の評価値を計算し提示することで電子掲示板内のスレッド選択及び閲覧を支援するシステムの構築を目指す。

2. 関連研究

2.1 評判情報の抽出に関する研究

電子掲示板から評判表現と表現情報の抽出を行う研究がある [1]。この研究では、電子掲示板の書き込みから統計的に評価表現を取り出し、肯定・否定の意味合いがどの程度強いかで評価表現をスコアリングし、出現評価表現の評価値の総和で評判情報の抽出を行う手法を提案している。評価表現には、形容詞、形容動詞の属性を使用している。また、Weblog を対象とした評価表現の抽出を行う研究がある [2]。本研究では、使用する単語に形容詞、形容動詞だけでなく動詞、名詞の属性および命令形の品詞を使用する点で前者と異なり、文の評価に係り受けは使用せず単語のみを使用し評価する点で後者と異なる。

連絡先: 砂山渡, 広島市立大学大学院情報科学研究科システム工学専攻, 広島市安佐南区大塚東三丁目 4 番 1 号, sunayama@sys.info.hiroshima-cu.ac.jp



図 1: 雰囲気を同定するシステムの構成

2.2 雰囲気の同定に関する研究

電子掲示板内のコミュニケーションを阻害する行為(荒らし行為)を単語辞書を用いて発見する研究がある [3]。この研究では、単語辞書を用いてレスを評価し得点付けを行い視覚化する事で、雰囲気を悪くする単語を含むレスが多く出現する部分に荒らし行為を発見する手法を提案している。また、番組実況チャットを利用したテレビ番組のメタデータ自動抽出を行う研究がある [4]。本研究では、雰囲気の良い部分だけではなく、雰囲気の良い部分の同定を行い閲覧者の雰囲気感覚に近づける事で閲覧支援を行う点で前者と異なり、喜びや落胆だけでなく、共感するレスや悪口で互いを揶揄するレスの評価を行う事で、雰囲気の良い悪いの同定を行っている点で後者と異なる。

3. 電子掲示板内の雰囲気を同定するシステム

本章では、電子掲示板内の雰囲気を同定するシステムについて述べる。提案システムの全体構成を図 1 に示す。提案システムは、スレッドの html ファイルを入力とする。入力となるスレッドの html ファイルからレス(書き込み)情報を抽出し、各レスの雰囲気の良い悪いの評価を行う。レスの評価後、レスの評価値を用いてレス間の関係に基づくレスの評価値を計算しスレッド内の雰囲気の変化の判定を行う。レスの評価とスレッド内の雰囲気の変化を雰囲気情報としてインタフェース上に出力を行う。

3.1 入力: スレッドの html ファイル

入力は、電子掲示板のスレッドの html ファイルとする。対象とするスレッドの html ファイルは、50 レス以上のスレッドの html ファイルとする。

3.2 レスの評価

本節では、抽出されたレス情報から単語特性辞書を用いてレスの評価を行う方法について述べる。ユーザが電子掲示板に書き込んだレスの内容を「茶釜」[5]を用いて形態素解析したのに対して単語特性辞書とのマッチングを行う事によりレスの評価を行う。

3.2.1 単語特性辞書

単語特性辞書は、閲覧者を不快にさせる単語を攻撃的な単語集合 (*Off*) と閲覧者の気分を良くする単語を回復的な単語集合 (*Def*) の2つの単語集合からなる。各集合の単語は式(1)、式(2)のカテゴリに属する単語から構成され、悪口は侮蔑、侮辱及び蔑称となる単語、ブラックは閲覧者に嫌な物(事)をイメージさせる単語、幸福は閲覧者に良い事をイメージさせる単語とする。攻撃的な単語集合の単語数は355と命令形の単語数の合計数とし、回復的な単語集合の単語数は276とする。これらの単語は「2ちゃんねる」内のスレッドを3000程度用いて頻度の高い単語を中心に集め辞書の作成を行った。

$$Off = \{ \text{悪口, 性風俗用語, ブラック, 命令形の単語} \} \quad (1)$$

$$Def = \{ \text{褒め言葉, 幸福, 挨拶言葉} \} \quad (2)$$

3.2.2 レスの雰囲気の良い悪しの評価

レスの評価値 Mp_i を用いてレスの良し悪しの評価を行う。 i 番目のレスの評価値 Mp_i を式(3)で定義する。 i 番目のレスの評価値 Mp_i は、レス内での回復的な単語集合の出現回数と攻撃的な単語集合の出現回数の差により計算する。レスの評価値 Mp_i が正であれば雰囲気の良いレスと評価し、負であれば雰囲気の悪いレスと評価する。

$$Mp_i = \sum_{def \in Def} def - \sum_{off \in Off} off \quad (3)$$

3.3 スレッド内の雰囲気の変化を判定

本節では、スレッド内の雰囲気の変化の判定方法について述べる。レス間の関係に基づきレスの評価を行い、雰囲気の流れの評価を行う事でスレッド内の雰囲気の変化の判定を行う。

3.3.1 レス間の関係に基づくレスの評価

レス間の関係に基づく i 番目のレスの評価値 Rp_i を式(4)で定義する。 Anc_i は、 i 番目のレス内におけるアンカーの有無を表す値とする。アンカーとは、ある特定のレスへの発言や返信を明確に示す記述の事である。アンカーが有れば2とし、無ければ1とする。 $Bonus_i$ は、 i 番目のレスの時点での雰囲気の良い悪しの評価それぞれの連続回数を表す値とする。

$$Rp_i = \begin{cases} -\log((|Mp_i| \times Anc_i + 1) \times Bonus_i) & (Mp_i < 0) \\ \log((Mp_i \times Anc_i + 1) \times Bonus_i) & (Mp_i \geq 0) \end{cases} \quad (4)$$

3.3.2 雰囲気の流れの評価

i 番目のレスを読む時点での雰囲気の流れの評価値 $A(Rp_i)$ を式(5)で定義する。 $A(Rp_i)$ が正であれば、 i 番目のレスを読む時点では雰囲気の良い流れであると評価し、負であれば、雰囲気の良い流れであると評価する。また、 W_k は i 番目を含む過去50レスそれぞれに対する重みとし式(6)で定義する。過去のレスほど重みを軽くする。これは、 i 番目のレスの時点の雰囲気は、 i 番目のレスが一番影響を与え、 i 番目から過去のレスほど印象が薄くなり雰囲気に関与しなくなると考えたか



図2: 雰囲気を同定するシステムの出力例

らである。また、過去50レスを用いる事については、ユーザが雰囲気を考える際に過去50レス程度のレスの内容を意識すると考えたからである。

$$A(Rp_i) = \sum_{k=0}^{50} Rp_{i-k} \times W_k \quad (5)$$

$$W_k = \frac{50 - k}{50} \quad (6)$$

3.4 出力: スレッドの雰囲気の流れの表示

本節では、システムの出力であるスレッドの雰囲気の流れについて述べる。雰囲気を同定するシステムの出力例を図2に示す。出力は、インタフェース上に「スレッド内の雰囲気の流れの良し悪しのグラフ」「レス内容および各レスごとの雰囲気の良い悪し」「タイトルと全体の雰囲気の良い悪し」「スレッド一覧」の表示を行う。「スレッド内の雰囲気の流れの良し悪しのグラフ」は、グラフの値(雰囲気の流れの評価値)が正であれば雰囲気の良い流れを表しており、負であれば雰囲気の良い流れを表している。「レス内容および各レスごとの雰囲気の良い悪し」は、レス内容の表示を行い、それぞれのレス番号の左に表示されている2つの顔のマークにおいて右側の顔マークでレスの雰囲気の良い悪しの評価表示を行っている。レスの雰囲気が良ければ笑顔のマーク、レスの雰囲気が悪ければ怒った顔のマーク、どちらでもなければ無関心の顔のマークを表示する。以降、顔のマークの表示方法は上記と同様である。さらに左側の顔のマークでは、レスを読む時点での雰囲気の流れの評価表示を行っている。「タイトルと全体の雰囲気の良い悪し」は、スレッドのタイトルと雰囲気の流れの評価値の全体平均を用いて全体の雰囲気の良い悪しを顔マークで表示している。「スレッド一覧」は、入力した複数のスレッドのhtmlファイルの情報をテーブルで表示する。

4. 雰囲気を同定するシステムの検証実験

本章では、提案する雰囲気を同定するシステムが、スレッド内の雰囲気をどの程度正しく同定できるかを検証した実験について述べる。

4.1 実験方法

被験者は普段からインターネットをよく利用する20代前半の男女12名であった。被験者1人あたりに8つのスレッド(1スレッドあたり連続する100レス)を読んでもらい、雰囲気

表 1: 実験に用いたスレッド

No.	スレッドタイトル
1	海外ドラマの魅力
2	最近の公務員叩きは異常 優秀なキャリア官僚のおかげで日本は裕福でいられる
3	おせちの不味さは異常 正月にあんなものを食べなければならぬかと思うと吐き気がする
4	読書しない奴ってなんなの？ 感性も知識もスラスカ 正直人間に産まれた意味ないよ
5	納豆の美味しい食べ方
6	若者の日本酒離れが止まらない 若者「無理して飲んでたけど、別に美味しくないからいらない」
7	【酒】 お前も年末年始用の良いお酒を買ったよな？
8	延命に年間4100万かかる男が正負に援助を要求 「政府の援助が無ければ死んでしまう。死ねと言うのか」
9	そろそろ彼女作れよ… いらぬ理由並べても虚しいぞ
10	AKB48 で一番かわいいのって
11	一生独身の魅力 結婚しちゃう情弱の皆さん、 俺たち独身で楽に生きててサーセン www
12	ガチで人生がつまらない奴こいよ

良いレスと雰囲気の良いレスの選択および雰囲気の良いレス区間と雰囲気の良いレス区間を選択してもらった。以下に1スレッドの作業のステップを示す。

- 100 レスを読み、雰囲気の良いレスと雰囲気の良いレスを選択する
- 1 で選択したレスを参考にして、雰囲気の良いレス区間と雰囲気の良いレス区間を選択する

実験では、「2ちゃんねる」内の12のスレッドを用意し、1スレッドを被験者8人に評価してもらった。使用したスレッドを表1に示す。

4.2 評価方法

実験により得られるデータは、被験者が選択した各スレッドのレスの雰囲気の良い悪いの評価およびレスの雰囲気の良い悪いの区間の評価である。被験者が選択した各スレッドのレスの雰囲気の良い悪いと提案システムおよび比較システムが同定した各スレッドのレスの雰囲気の良い悪いの適合率、再現率、F値を求め、また、レスの雰囲気の良い悪いの区間と提案システムおよび比較システムが同定した各スレッドのレスの雰囲気の良い悪いの区間のグラフを比較する事で、雰囲気を同定するシステムを評価する。

4.3 比較システムの詳細

比較システムには、提案した雰囲気を同定するシステムの単語特性辞書をスピンモデルによる単語の感情極性抽出 [6] で作成される P/N 単語集合に入れ替えたものを用いた。P/N 単語集合とは、一般的に良い印象を持つ単語をポジティブ (P)、一般的に悪い印象を持つ単語をネガティブ (N) とし、-1 から +1 の実数値を割り当てた単語集合の事であり、-1 に近いほどネガティブな単語を表し、+1 に近いほどポジティブな単語を表す。単語特性辞書の攻撃的な単語集合に一般的に悪い印象を持つネガティブ (N) の単語を使用し、回復的な単語集合に一般的に良い印象を持つポジティブ (P) の単語を使用した。また、今回は、ポジティブ (P) とネガティブ (N) の意味合いをある程度強く持つ単語を用いる事にしたため、ポジティブ (P) には実数値 0.99 以上の単語選択し、ネガティブ (N) には実数値 -0.99 以下の単語を選択した。攻撃的 (ネガティブ) な単語集合の単語数 3287 はであり、回復的 (ポジティブ) な単語集合の単語数 1058 はである。

表 2: レスの評価に対する適合率、再現率、F 値

	提案システム		比較システム	
	雰囲気 (良い)	雰囲気 (悪い)	雰囲気 (良い)	雰囲気 (悪い)
適合率	0.72	0.84	0.64	0.64
再現率	0.59	0.60	0.18	0.68
F 値	0.64	0.69	0.26	0.60

表 3: 1 スレッドあたりの雰囲気の良い (良く) するレスの数

	被験者	提案	比較
雰囲気の良いレス数	20.9	14.1	20.2
雰囲気の良いレス数	17.3	13.5	4.8

4.4 実験結果と考察

被験者が選択した各スレッドのレスの雰囲気の良い悪いとシステムの出力結果を比較した。提案システムと比較システムそれぞれの雰囲気の良いレスおよび雰囲気の良いレスに対する適合率、再現率、F 値を求めた結果を表2に示す。表2を見ると、提案システムの雰囲気 (悪い) の再現率を除く適合率、再現率ともに比較システムを上回っていた。これは、提案システムの方が高い精度でレスの雰囲気の良い悪いの評価をしていることを示している。提案システムの雰囲気 (悪い) の再現率が比較システムの雰囲気 (悪い) より低いことに関しては、表3を見ると、提案システムの雰囲気の良い (良い) レス数は同程度であり被験者選択に対して雰囲気の良い (良い) のどちらとも約7割から8割のレスの抽出を行っている。しかし、比較システムの雰囲気の良い (良い) レス数は抽出数に大きな差が見られ、被験者選択に対して雰囲気の良いレスはほぼ同程度の抽出数であったが、雰囲気の良いレスは約3割程度の抽出数と低い値を示していた。原因は、単語辞書の単語数に違いがあるためだと考えられる。その理由は表4を見ると、提案システムは、攻撃的な単語数と回復的な単語数がほぼ同程度の数であったのに比べ、比較システムは攻撃的な単語数と回復的な単語数の比が約3:1となり大きな差があった。この差により、比較システムでは、攻撃的な単語集合を回復的な単語集合より多く抽出され評価が行われたため、雰囲気の良いレス数が多くなり、雰囲気の良いレス数が少なくなったのだと考えられる。また、これらの結果より、レスの評価を行う際に使用する単語については、単語数が多いよりも選別され必要であるカテゴリーと単語からなる単語特性辞書を作成し用いる事が効果的だとわかる。

被験者が選択した各スレッドのレス区間の雰囲気の良い悪いとシステムの出力結果を比較した。表1のスレッド4において提案システムおよび比較システムの区間評価 (雰囲気の流れの評価) と被験者の区間評価平均をグラフにしたものを図3に示す。横軸がレス番号、縦軸が雰囲気の良い悪いの区間評価値を表す。

図3を見ると、提案システムの区間評価と被験者の区間評価平均は所々違いは見られるもののおおよそグラフが一致している事が見てとれた。また、比較システムの区間評価と被験者の区間評価平均はあまり一致しておらず、グラフも雰囲気の良い悪い (雰囲気の良い悪い流れ) が始終続いていた。これは、提案システムが雰囲気の良い悪いの流れ (変化) をおおよそ同定できている事を示しており、提案システムの雰囲気の良い悪いの流れ (変化) の同定の精度が比較システムの雰囲気の良い悪いの流れ (変化) の同定より高い事を示している。比較システムの区間評価が始終雰囲気の良い悪い区間 (雰囲気の良い悪い流れ) だっ

表 4: 提案システムと比較システムの攻撃的および回復的な単語数

	提案システム	比較システム
攻撃的な単語数	355	3287
回復的な単語数	276	1058
合計単語数	631	4345

た事に関しては、上記のレス評価でも示したとおり、レスの評価を行う時点で雰囲気の良いレスの評価が多く雰囲気の悪いレスの評価が少なかったため、区間評価においても、終始雰囲気の悪い状態が続いてしまったと考えられる。図3について、30から45レス付近で被験者の区間評価平均と大きな差が見られる事に関しては、提案システムでのレス評価において、「だる」、「してやってもいいけど」と上から目線でのレスや相手を煽るレス、「言葉知らんの?」、「ばくちゃんご本よんでお勉強してます」と相手を馬鹿にするレスを雰囲気の悪いレスと評価できていなかった。これらは、単語だけでは抽出する事が出来ないため、本システムでは、レスの雰囲気の評価は出来ない。しかし、構文を解析する事で抽出が可能となり同定ができると考えられる。これは今後の課題となる。

全体的に雰囲気が良くなり始める点、雰囲気が悪くなり始める点について、被験者の区間評価平均では激しく上昇や下降しているのに対して、提案システムの区間評価ではなだらかに上昇や下降するので被験者の感覚に対して遅延が発生している事が見てとれた。この原因としては、提案システムは区間の評価(雰囲気の流れの評価)をする際に過去50レスを用いて最新のレス程重視する重み付けを行い計算しているが、雰囲気が良く(悪く)なり始める1レスの評価値では急激に区間の評価値(雰囲気の流れの評価値)を変化させる程の値にはならないためと考えられる。これは、過去50レスを用いる事の検討、各レスに対する重み付けの検討、区間の始まりと区間の終わりを抽出する事で区間を形成する区間変換アルゴリズムの検討を行う事で、さらに、被験者の雰囲気の変化に近づける事が可能になると考えられる。

比較システムの区間評価において終始雰囲気の悪い区間(雰囲気の悪い流れ)が続いた事やほとんど回復に転じなかった事について、上記で述べたとおり、雰囲気の悪いレスの評価が多く雰囲気の悪いレスの評価が少ない事が問題である事がわかる。その原因は、比較システムの単語辞書(P/N単語集合)の単語数が多く、攻撃的な単語数(N:ネガティブ)と回復的な単語数(P:ポジティブ)の比率が3:1と大きく差が出てしまった事が考えられる。しかし、提案システムでは、有効な単語を選別し、少ない単語数に抑え攻撃的な単語数と回復的な単語数の比率にあまり差を出さずに単語特性辞書を作成する事により精度の高いレスの評価を行う事ができたため、レスの区間の雰囲気の良い悪いの変化を同定する事ができたと考えられる。この結果からも、単語特性辞書の作成は、レスおよびレスの区間の雰囲気の良い悪いの評価に対して効果的である事がわかる。

これらの結果より、レスの区間評価には課題がまだあるものの、おおよそのレス区間の雰囲気の同定は行える事と、さらに精度の高いレス区間の雰囲気の同定の可能性を確認した。

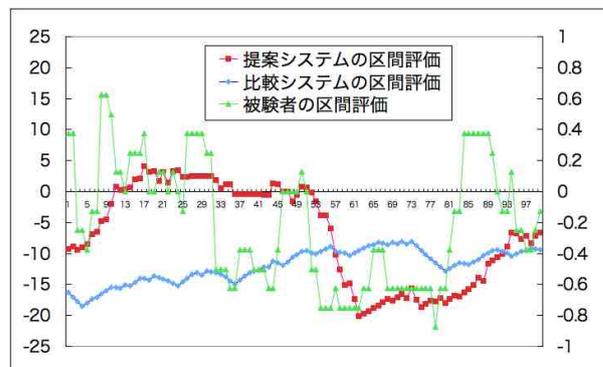


図 3: 提案システムおよび比較システムの区間評価と被験者の区間評価平均のグラフ (スレッド 4)

5. 結論

本研究では、電子掲示板内において雰囲気の良い部分と悪い部分を同定するシステムを提案した。

雰囲気を同定するシステムの検証実験により、提案システムが電子掲示板内のコミュニケーションにおける雰囲気の良い悪いおよび変化を同定していることを確認した。特に、レスの雰囲気の良い悪いを評価する際に用いる単語特性辞書の作成は、雰囲気の良い悪いを同定する上で効果的であることがわかった。

今後の課題としては、単語特性辞書の単語それぞれに適切な評価値をスコアリングする事により精度の高い雰囲気の同定を行う事を目標とする。

参考文献

- [1] 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, 3F1-03, (2004)
- [2] 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)
- [3] 一藤裕, 今野将, 曾根秀昭: テキストベースコミュニケーションにおける阻害行為に関する評価手法の提案, 情報処理学会研究報告, QAI, [高品質インターネット], Vol.106, No.62, pp.43-47, (2006)
- [4] 宮森恒, 中村聡史, 田中克己: 番組実況チャットを利用したテレビ番組のメタデータ自動抽出方式, 情報処理学会論文誌, データベース, Vol.46SIG18, No.28, pp.59-71, (2005)
- [5] 松本裕治, 山下達雄, 平野義隆, 松田寛, 高岡一馬, 浅原正幸: 形態素解析システム『茶筌』, Ver.2.4.0, 使用説明書, (2007)
- [6] 高村大也, 乾孝司, 奥村学: スピンモデルによる単語の感情極性抽出, 情報処理学会論文誌ジャーナル, Vol.47, No.2, pp.627-637, (2006)