

ネタバレと興味喚起を考慮した小説レビューのランキング手法

村井 聡[†] 牛尼 剛聡^{††}

[†]九州大学大学院芸術工学府 〒815-8540 福岡市南区塩原 4-9-1

^{††}九州大学大学院芸術工学研究院 〒815-8540 福岡市南区塩原 4-9-1

E-mail: [†]2ds11094e@s.kyushu-u.ac.jp, ^{††}ushiama@design.kyushu-u.ac.jp

あらまし 小説のレビューは、未読者がその小説を読むかどうかを判断する指標となる。既読者は、レビューに興味を感じた内容を書く際に「ネタバレ」(未読小説の読書時の楽しみを奪う記述)を同時に記述することがある。未読者にレビューを提示する際に「ネタバレ」を防ぐことは重要である。しかし「ネタバレ」と考えられる内容の中には、既読者が興味を持った内容が含まれており、単純な感想や、あらすじの紹介よりも未読者に対する訴求力が強いと考えられる。本研究では、レビューによって付加された「ネタバレ」を表すタグを用いて「ネタバレ」にならない範疇で読者の興味をひくレビューを優先的に提示することを目標とする。なお、本論文の第2章から第7章は「ロートレック荘事件」(著者:筒井康隆)の「ネタバレ」を含むため、読書中、もしくは今後読む予定がある方は注意して欲しい。キーワード レビュー,ネタバレ,興味喚起,小説

A Method for Ranking Reviews of A Novel Based on Their Spoiler and Attractiveness

Soichi MURAI[†] and Taketoshi USHIAMA^{††}

[†] Graduate School of Design, Kyushu University
4-9-1 Shiobaru, Minami-Ku, Fukuoka 815-8540 Japan

^{††} Faculty of Design, Kyushu University
4-9-1 Shiobaru, Minami-Ku, Fukuoka 815-8540 Japan

E-mail: [†]2ds11094e@s.kyushu-u.ac.jp, ^{††}ushiama@design.kyushu-u.ac.jp

1. はじめに

近年、iPad や Kindle などの電子書籍を閲覧できるデバイスの普及にともない、電子書籍市場は急成長している。また、Amazon をはじめとするオンライン書店の利用が一般化し、多くの人々が Web 上で書籍を選び、購入するようになった。Web 上で書籍を選択する機会の増加に伴い、ユーザにとって価値のある書籍を効率的に見つける機能が重要になっている。これまでに、大量の書籍の中から閲覧する候補を絞り込む事を目標として、検索・推薦に関する研究が行われてきた。しかし、これらのシステムを使用する際に、ユーザは検索や推薦された書籍をすべて読む訳ではない。ユーザは、検索された書籍や推薦された書籍の中から、自らの判断で読む書籍を選別することが一般的である。

ユーザが読んだことのない書籍が、自分にとって価値のあるものを判断する際に、既読者の感想であるレビューは選別の参考になる。小説の選別においては、レビューは読書時の楽

しみの期待度を推測する参考になる。本研究では、小説のレビューにおける尺度を「ネタバレ」と「興味喚起」の二点から考える。

小説のレビューにおいては、読書時の楽しみを損なわせる「ネタバレ」を含むものが存在する。レビューを書くモチベーションとしては「ネタバレ」をするためにレビューを書くのではなく、自分が興味を感じた内容を記録するため、もしくは他の人と共有するためにレビューを書くと考えられる。しかし、結果的に、未読者にとって「ネタバレ」と位置づけられる内容を含むことがあるそのため、「ネタバレ」を含んだレビューを、ユーザの同意無しに見せない機能を実現することは有意義である。

現在、SNS や web サービスを利用して、ユーザがレビューを気軽に投稿できるようになった。ユーザが「ネタバレ」を含む内容のレビューを投稿する際には、それが「ネタバレ」を含むことを明記することが行われている。しかし「ネタバレ」の定義は人それぞれであるため「ネタバレ」であると記述されてい

ても、一般的には「ネタバレ」ではないと考えられるレビューが存在する。逆に、「ネタバレ」と明記されていなくても、「ネタバレ」と考えられるレビューも存在する。多くのユーザはレビューを書くための特別な訓練を受けたわけではなく、「ネタバレ」かどうかの判断は主観的に行っている。

「ネタバレ」を含むレビューは、ユーザの読書の楽しみを損なう可能性がある。しかし、小説本文の内容に触れないレビューでは、小説の表面的な感想しかわからない。これまでに「ネタバレ」を防ぐ研究、もしくは、レビューの感情・意見の抽出は行われてきたが、「ネタバレ」を考慮に入れた意見抽出に関する研究はあまりなされていない。

本研究では、「ネタバレ」にならない範囲で、小説中で読者の興味をひく内容について記述しているレビューを優先的に見せることにより、読者の小説の選別を支援できると考え、レビューを「書籍の「ネタバレ」となる度合い」と「読者の興味喚起を喚起する度合い」の二つの観点から評価し、ランク付けにより推薦する。アプローチとしては、最初に単語の「ネタバレ」の重みを決定する。それを元に、各レビューを句読点で区切ったフレーズが「ネタバレ」に関する度合いを決定する。次に、各フレーズが未読者の興味をひく度合いを計算する。最後に、「ネタバレ」に基づいたランキングと「興味喚起」に基づいたランキングから、ユーザに提示する最終的なランキングを決定する。

本論文の構成は以下の通りである。第2章では、関連研究について述べる。第3章では、単語の「ネタバレ」に関する度合いを決定する。第4章では、「ネタバレ」に関する語から「ネタバレ」内容を含むフレーズを抽出する。第5章では、被験者実験により「ネタバレ」内容を含むフレーズを抽出することができたかを評価する。第6章では、「ネタバレ」内容を含むフレーズが未読者の興味をひく度合いを決定する。第7章では、「ネタバレ」にならない範囲で小説本文の内容にふれており、かつ、未読者の興味をひくフレーズのランキングを求める方法について述べる。第8章では、今後の課題について述べる。

2. 関連研究

本研究の目的と近い研究として、池田らの研究がある[2]。池田らは、あらすじ(ストーリーの内容)部分に出現しやすい語と、出現しにくい語があると考え、単語を素性とした機械学習により、本や映画などからストーリーの内容を見つけ出す手法を提案している。しかし、本研究では、ネタバレに関する語は書籍において異なると考える。例えば、「ロートレック荘事件」の「ネタバレ」には『車中では三人で会話していたことを隠すために、一人称を使って書かれている』ことに関する内容が含まれている。この「ネタバレ」に関する語は『会話』『一人称』といった、一般的には「ネタバレ」に関しそくない単語となる。本研究では、レビューがネタバレを含むかどうかを明示する「ネタバレタグ」を利用して、書籍毎に単語の「ネタバレ」に関する度合い(ローカルネタバレ度)を決定する。

中村[1]は「ネタバレ」を防止するためのフィルタリング手法を提案している。本研究では、小説のレビューをランキング

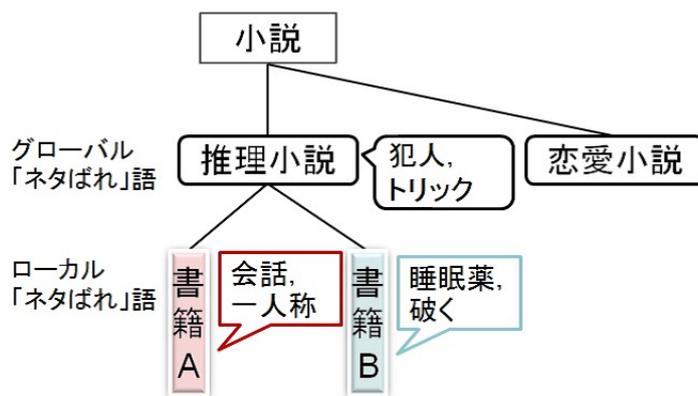


図1 「ローカルネタバレ語」と「グローバルネタバレ語」

で提示する。その際に「ネタバレ」を含むレビューを下位に表示することで「ネタバレ」を含むレビューが、ユーザの目に触れないようにする。

レビューのランキングに関する研究では、小倉ら[3]はレビューの品質を考慮したランキングを行なっている。本研究では、小説の選別において、レビューの信頼度の尺度の他に、未読者の興味を喚起させる尺度も重要だと仮定し、興味をひく度合いを考慮に入れたレビューのランキングを行う。

3. 「ネタバレ」に関する語の抽出と重み付け

本研究では、「ローカルネタバレ語」、「グローバルネタバレ語」の二種類のネタバレ語を抽出する(図1)。「グローバルネタバレ語」はその小説が属するジャンルにおいて「ネタバレ」に関連する単語であり、「ローカルネタバレ語」は、ある小説において「ネタバレ」に関連する単語である。「グローバルネタバレ語」の例としては、推理小説における「犯人」、「動機」等が考えられ、ある程度ヒューリスティックに決定することが出来る。一方「ローカルネタバレ語」は、小説ごとに異なるため、一冊ずつヒューリスティックに決定していくのは困難である。本研究では、「ローカルネタバレ語」の推定手法を提案する。本研究では、多くの既読者が「ネタバレ」であると考えた記述に頻出する単語は「ネタバレ」に関する単語である可能性が高いと考え、Web上にある「ネタバレ」として明示的に示されたテキストを利用する。何を「ネタバレ」であるか考えるかは、レビューによって異なるため、多くのレビューが「ネタバレ」であると明記した文章に基づいて、「ネタバレ」に関する度合いを決定する。

本研究では、「ネタバレ」が明記された文章として booklog(<http://booklog.jp/>) に投稿されたレビューを利用する。booklogでは、2010年12月3日にレビューが「ネタバレ」であることを明示するタグを付ける機能が追加された。追加された機能が広まるまで、ある程度の時間がかかると考え、2011年1月1日以降のレビューを対象とした。また、小説のジャンルによって、ネタバレに関する語は異なると考えられるため、ジャンル毎に「ネタバレ」に関する重みを決定していくことで精度が上がると思われる。本稿では予備実験として推理小説を用いた。用いた小説のタイトルは「推理小説タグ」が付いた799

タイトルのうち「ネタバレ」タグ(有)、「ネタバレ」タグ(無)の両方のレビューが10以上存在する117タイトルである。

3.1 ローカルネタバレ度

単語 t_i が、ある小説 B_j において、ネタバレに関する可能性をネタバレタグが付いたレビューに出現する度合いに基づいて、ローカル「ネタバレ」度を決定する。ローカル「ネタバレ」度 Spoiler-Local(t_i, B_j) を以下の式で決定する。

$$\text{Spoiler-Local}(t_i, B_j) = \text{SpoilerTag-Weight}(t_i, B_j) \cdot \text{idf}(t_i) \quad (1)$$

ここで、 $\text{SpoilerTag-Weight}(t_i, B_j)$ は単語 t_i が小説 B_j の「ネタバレ」タグ(有)のレビューに出現する度合いを表し、以下の式で定義される。

$$\text{SpoilerTag-Weight}(t_i, B_j) = \frac{n_{t_i, B_j, \text{tag}=1}}{n_{t_i, B_j, \text{tag}=1} + n_{t_i, B_j, \text{tag}=0} \cdot \frac{|\{r_{\text{tag}=0} : r \in R(b), r \ni t_i\}|}{|\{r_{\text{tag}=1} : r \in R(b), r \ni t_i\}|}} \quad (2)$$

また、 $\text{idf}(t_i)$ は逆文章頻度であり、ある小説のレビュー集合を d とし、以下の式で定義する。

$$\text{idf}(t_i) = \log \frac{D}{|\{d : d \ni t_i\}|} \quad (3)$$

ここで $n_{t_i, B_j, \text{tag}=1}$ は、小説 B_j において「ネタバレタグ」が付いたレビューに出現する t_i の数を表し、 $n_{t_i, B_j, \text{tag}=0}$ は、小説 B_j において「ネタバレタグ」が付いていないレビューに出現する t_i の数を表す。また、 $|\{r_{\text{tag}=1} : r \in R(b), r \ni t_i\}|$ は、小説 B_j において t_i を含んだ「ネタバレ」タグの付いたレビュー数を表す。ただし、 $n_{t_i, B_j, \text{tag}=0} + n_{t_i, B_j, \text{tag}=1} = 1$ (その小説のレビューに一度しか出現しない単語)の時は、 $\text{Spoiler-Local}(t_i, B_j) = 0$ とする。

ここで、対象とする単語は、名詞(代名詞、接尾、数を除く)、動詞(自立動詞以外を除く)、形容詞である「する」、「ある」、「よる」、「いる」、「なる」、「いう」、「みる」、「できる」の8形態素は出現頻度が高いが、他の語句の補助的な機能を持つ語句であることから、興味の強弱の判断材料として不向きであるため、ストップワードと設定した[4]。形態素に分ける際に、平仮名一文字、カタカナ一文字もストップワードとした。形態素解析には、オープンソースの形態素解析エンジン MeCab[5]を用いた。

「ロートレック荘事件」において高い Spoiler-Local(t_i, B_j) をとる単語(ただし SpoilerTag-Weight(t_i, B_j) > 0.75)を表1に示す。

前記した「ネタバレ」に関する『一人称』、『フルネーム』といった単語も上位に出現する『叙述』(「叙述トリック」)を「ネタバレ」に関する単語と捉えるかは、ユーザによって異なると思われる。

本研究では、以下の理由から単語単位での「ネタバレ」に関するかどうかの評価は行わない。未読小説の「ネタバレ」の内容を、単語単位で評価するのは難しいからである。また、本研究の目的は、レビューの推薦であり、単語と単語がつながったレビューとして提示された時に「ネタバレ」であるかないかが重要であるからである。

表1 「ロートレック荘事件」において高いローカル「ネタバレ」度をとる単語

単語	Spoiler-Local
フルネーム	4.069
見取り図	3.376
注釈	3.052
面倒	1.584
一人称	1.504
ルール	1.361
叙述	1.021
飛ばす	0.978

3.2 グローバルネタバレ語

あるジャンルの「ネタバレ」を記述する際に、出現しやすい単語がある。本論文ではこれらの単語を、グローバル「ネタバレ」語と呼ぶ。推理小説のグローバル「ネタバレ」語としては、「トリック」、「犯人」、「動機」等が考えられる。これらの単語は、ある小説のレビューにおいて SpoilerTag-Weight(t_i, B_j) が低くなったとしても、たまたま値が小さくなったと考え、「ネタバレ」に関する語として扱うのが妥当だと考えられる。

グローバル「ネタバレ」語を含むレビューには、「トリックが面白い」、「トリックが難しい」といった、語に対する評価の記述も多く含まれている。グローバル「ネタバレ」語を含み、かつ、その小説特有の単語を含んでいるレビューは、ネタバレに関するレビューだと考えられる。グローバル「ネタバレ」語の推定は今後の課題とする。

4. フレーズの「ネタバレ」度

「ネタバレタグ」の有無に関わらず「ネタバレ」に関する内容と「ネタバレ」に関しない内容の両方を含むレビューが存在する。そこで、まず、レビューを句点により、分けたフレーズに分ける。そして、それぞれのフレーズに対して「ネタバレ」に関する度合いを決定する。今回は、 $\text{SpoilerTag-Weight}(t_i, B_j) > 0.75$ となる単語を「ネタバレに関する語」とし、各フレーズの「ネタバレ」の度合いを決定する。

しかし、これだけでは「ネタバレに関する語」が少ないので、「ネタバレに関する語」の拡張を行う。「ネタバレに関する語」と同フレーズに出現する単語は、「ネタバレ」に関する語である可能性が高いとし、対象小説のレビューにおいて、二回以上同じフレーズに出現し、 $\text{idf}(t_i) > 0$ となる単語も「ネタバレに関する語」として扱う。ただし、小説紹介やあらすじ等においても、同フレーズに出現する単語の組み合わせは、事前に提示されている小説の内容であり、ネタバレとなる可能性が低いと考え、「ネタバレに関する語」から除外した。

フレーズ p_i のネタバレ度 Spoiler(p_i, B_j) は、「ネタバレに関する語」のネタバレ度を足し合わせることで決定する。

$$\text{Spoiler}(p_i, B_j) = \sum_{t_i \in T(p_i)} (\text{Spoiler-Local}(t_i, B_j)) \quad (4)$$

ただし、フレーズは単語と単語との繋がりの中でできると仮定し、「ネタバレに関する語」を一つしか含まない Spoiler(p_i) の

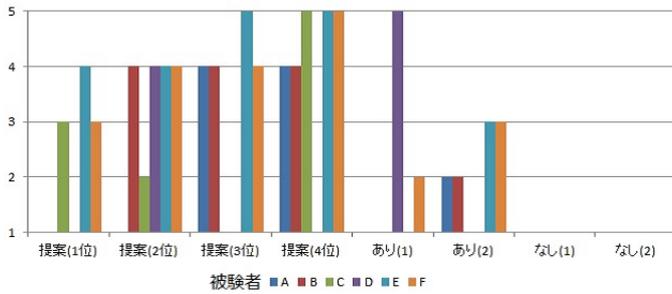


図 2 各フレーズに対する，各被験者の評価

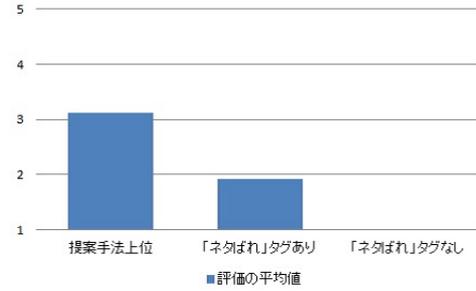


図 4 各種フレーズに対する，被験者の評価の平均値

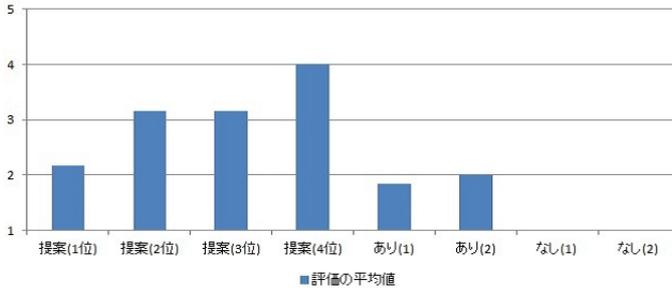


図 3 各フレーズに対する，被験者の評価の平均値

値は 0 とする。

「ロートレック荘事件」において，高い $Spoiler(p_i)$ をとる上位 4 フレーズを以下に示す． $Spoiler(p_i)$ の計算に利用した単語を太字で表記する。

(1) 名字だけ・名前だけだったり，フルネームだったり，一人称で誰のセリフなのかよく分かりづらかったり。

(2) 気づこうと思えば気づいてしまう見取り図の使い方が悪い。

(3) おかしいなと思い始めたのは登場人物紹介のページがなかったことと，見取り図を見たとき。

(4) どうして一人だけフルネーム？と思ってたら別々の人物で二人いたからだった。

5. 予備実験

5.1 実験内容

提案手法により「ネタバレ」となるフレーズを抽出することができたかを，被験者実験により評価する．評価対象の小説としては「ネタバレ」が明確である「ロートレック荘事件」を選択した。

被験者には「商品説明(あらすじ)」を読んだ後に，提示されたフレーズが「読書時の楽しみを奪うと思うか」について 5 段階で評価してもらった(5: 思う, 4 どちらかと言えば思う, 3: わからない, 2: どちらかと言えば思わない, 1: 思わない)．提示したフレーズは，高い $Spoiler(p_i)$ をとる上位 4 フレーズ，「ネタバレ」タグ有りのレビューからランダムに選択した 2 フレーズ，「ネタバレ」タグ無しのレビューからランダムに選択した 2 フレーズの計 8 フレーズである．被験者数は 6 名であり，全員が「ロートレック荘事件」を読んだことがない。

5.2 結果と考察

提示された各フレーズに対する，被験者の評価を，図 2，図 3，図 4 となる．図 4 より，提案手法により「ネタバレ」と思われるレビューを抽出することができた．また，図 2 より，何を「ネタバレ」ととらえるかは，個人間のばらつきがあることもわかる。

図 3 より，提案手法の中で，一番高い評価の平均値を得た提案(4位)と，一番低い評価の平均値を得た提案(1位)に対して考察する．被験者に提示されたそれぞれのフレーズは以下のとおりである。

(1) 提案(1位): 名字だけ・名前だけだったり，フルネームだったり，一人称で誰のセリフなのかよく分かりづらかったり。

(2) 提案(4位) どうして一人だけ？と思ってたら別々の人物で二人いたからだった。

両フレーズとも「ネタバレ」(車中では三人で会話していたことを隠すために，一人称を使って書かれている)に関する記述となっている．しかし，提案(1位)のフレーズは，小説を読んだことがない被験者には「ネタバレ」に関する話題だと捉えられず，文章がわかり難かったという感想と捉えられた，と考えられる．一方，提案(4位)のフレーズが「ネタバレ」だと捉えられた理由としては「問いに答えている」「～からだった」という表現が使われていることが考えられる．未読者がそのフレーズが「ネタバレ」かどうかを判断する基準について，調査実験を行い，適切なレビューの提示方法を考えることを今後の課題とする．また，今回提案した手法では，ローカル「ネタバレ」語を推定するのに「ネタバレ」を含むかどうかを明記した記述が複数必要であったが「～からだった」等の表現を手掛かりに「ネタバレ」に関する語を推定していく手法についても，今後考えていく。

6. フレーズの興味喚起度

我々は，著作権の切れた小説を対象とし，小説本文とレビューを用いて，小説において興味をひく単語を抽出する手法を提案してきた [6]．具体的には，小説本文とそのレビューに出現する単語において，その小説のレビューに出現する頻度が高く，他の小説のレビューに出現する頻度の低い単語を，興味をひく度合いの高い単語として抽出した．書籍集合を B ，小説 b に対するレビューの集合を $R(b)$ とすると，小説とレビューの関係は

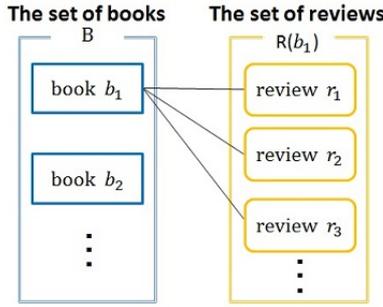


図5 小説とレビューの関係

図5のように表すことができる。注目する小説を b としたとき、単語 t_i の興味喚起度 at を以下の式で定義される。

$$at(t_i, b) = \sum_{r \in R(b)} tf(t_i, r) \cdot \log \frac{|B|}{|\{b : r \in R(b), r \ni t_i\}|} \quad (5)$$

本研究では、小説本文のデータが利用できない小説を多く扱うため、小説本文によるレビューのフィルタリングは行わない。 p_i の興味喚起度 $Attract(p_i)$ は、名詞、動詞、形容詞の $at(t_i, b)$ を足し合わせることで決定する。また、フレーズに出現する全ての単語が、あらずじ等に含まれている場合は、引用とみなし、ランキングからは除外する。

$$Attract(p_i, b) = \sum_{t_j \in T(p_i)} \frac{at(t_j, b)}{\sum_{i=k}^{n_{t_k}} n_{t_k}} \quad (6)$$

ここで $\sum_{i=k}^{n_{t_k}} n_{t_k}$ は、 p_i に含まれる単語数を意味する。「ロートレック荘事件」において、高い $Attract(p_i)$ をとる上位4フレーズを以下に示す。 $Attract(p_i)$ の計算に利用された単語を太字で表記する。

- (1) どうして一人だけフルネーム？と思ってたら別々の人物で二人いたからだった。
- (2) 気持ち悪さはこれだったのか！とすっきりしました。
- (3) 何が起こったかを当ててみる！
- (4) 最後まで真相がわからない傑作ミステリーでした。

7. フレーズのランキング

「ネタバレ」にならない範囲で、未読者の興味をひくフレーズのランキングを行う。 $Spoiler(p_i)$ の順位は低いほどネタバレに関する可能性が低く、また $Attract(p_i, b)$ の順位は高いほど未読者の興味をひく可能性が高い。それぞれの順位を $Rank_{Spoiler}$ 、 $Rank_{Attract}$ とし、フレーズの得点 $Score(p_i, b)$ を以下の式で求める (例：図6)。

$$Score(p_i, b) = \alpha \cdot Rank_{Spoiler(p_i, b)} + (1 - \alpha) \cdot (1 + \max(Rank_{Attract(p_i, b)}) - Rank_{Attract(p_i, b)}) \quad (7)$$

ユーザには $Score_{p_i, b}$ が高い順にフレーズを並び替えたものをランキングとして提示する。 $\alpha = 0.5$ とした時に、「ロートレック荘事件」において、高い $Score_{p_i, b}$ をとる上位4フレーズを以下に示す。なお、フレーズ冒頭の(ネ: 興:)は、「ネタバレ」でのランキング順位と、興味喚起でのランキングの順位を表す。

$Rank_{Spoiler}$	$Spoiler(p_i, b)$	$Rank_{Attract}$	$Attract(p_i, b)$
1	10(i=1)	1	10(i=3)
2	8(i=2)	2	9(i=1)
3	5(i=3)	3	8(i=2)
4	3(i=4)	4	7(i=4)
5	2(i=5)	5	2(i=5)

$$\text{例: } Score(p_1, b) = 1 + (5 - 2) = 4$$

図6 $Rank_{Spoiler}$ の求め方

(1) (ネ:45 興:2) 気持ち悪さはこれだったのか！とすっきりしました。

(2) (ネ:40 興:9) 第十七章からの「解」ではあまりにも丁寧にトリックが書かれて、ああ、そういうことか と納得。

(3) (ネ:39 興:11) ひどいネタバレになるからここには書かないけど、個人的にどうしても納得できない、気になってるところがあるので 四つで。

(4) (ネ:37 興:10) 実は車のなかに三人いたのか...車中では三人で会話してたのか...とやっぱりな展開でした。

結果をみると、小説の内容とはあまり関係のないフレーズが上位にランク付けされている。 $Attract(p_i, b)$ が、小説の内容とはあまり関係のないフレーズに対しても大きくなってしまっていること原因だと考えられる。今後小説の内容との関連を求める方法を改善する必要がある。また、「ネタバレ」に強く関連すると思われるフレーズが、4位にランキングしている「ネタバレ」と興味喚起の影響のバランスを最適化させる、 α の値を決定する手法の提案を今後の課題とする。

8. まとめ

本研究では「ネタバレ」にならない範囲で、小説の内容にふれている小説のレビューを抽出する手法を提案した。アプローチとしては、まず「ネタバレ」であるか否かを明示する「ネタバレタグ」に基づいて、「その小説に特有のネタバレ」となる単語を求め、次にフレーズの「ネタバレ」度を決定した。予備実験の結果、「ネタバレ」に関するフレーズの抽出ができた。しかし、フレーズが「ネタバレ」に関する単語を含んでいても、未読者が「ネタバレ」だと判断できない場合があることがわかった。今後は、未読者が「ネタバレ」と判断するレビューに構文的特徴があるか、また、レビュワーが「ネタバレ」として書くレビューに構文的特徴があるかについて調査する。他の課題としては、小説の内容に関するレビューを抽出する点で改善があげられる。この問題は、池田らの研究[2]を参考にし、小説の内容に関係ないフレーズを除去することで改善できると考えられる。

また、今後は小説が「ネタバレ」を持つ度合いの推定も行なっていく。本論文で例に上げた小説(ロートレック荘事件)は、はっきりとした「ネタバレ」があるが、そうでない小説もある。はっきりとした「ネタバレ」のない小説のレビューでは、「ネタバレ」に関する内容かどうかを判断することは難しい。例えば、「ドミノ」(著者:恩田陸)のレビューには「ネタバレ」と思わ

れるフレーズがない。そのような小説では、「ネタバレ」の重みよりも、「興味喚起」の重みを強くしたランキングを提示することで、ユーザの満足度を上げることができると考えられる。

文 献

- [1] 中村 聡史, "ネタバレ防止ブラウザの実現", WISS2010
- [2] 池田 郁, 土方 嘉徳, 西田 正吾, "レビューからのあらずじ除去における機械学習アルゴリズムの適用", WebDB Forum2010
- [3] 小倉達矢, 穴戸 開, " レビューサイトにおける良質なレビューの特性とそれを考慮した. 評判情報の抽出に関する一考察 ", DEWS2008.
- [4] 沢井康孝, 山本和英, " 文書に対する大衆の興味の強さの推定 ", 自然言語処理= Journal of natural language processing Vol.15 , No.2 , pp.101-136 , 2008-04-10
- [5] MeCab: Yet Another Part-of-Speech and Morphological Analyzer , <http://mecab.sourceforge.net/>.
- [6] Soichi Murai and Taketoshi Ushiyama "A Method for Extracting Attractive Sentences from an Electronic Book Based on Reviews for Effective Browsing", presented at 15th International Conference on Knowledge Based and Intelligent Information & Engineering Systems , 12 , 13 & 14 2011