

読者コメント連動型漫画における人物の登場頻度分析

An frequency analysis of characters in reader comments of online manga

本田優也¹ 廣川佐千男^{1,2} 伊東栄典²

Yuya Honda¹, Sachio Hirokawa^{1,2} and Eisuke Ito²

¹九州大学大学院ライブラリサイエンス専攻

¹ Graduate School of Library Science, Kyushu University

²九州大学情報基盤研究開発センター

² Research Institute for Information Technology, Kyushu University

Abstract: 現在漫画コンテンツの電子化が進んでいる。従来漫画の表現は、紙で印刷するものに適して進化してきた。スマートフォンの普及により、電子漫画の表現も変化している。電子コンテンツの場合、利用者（読者、視聴者）と作者との交流が、簡単かつ短時間でやりとり可能になっている。読者のコメントから、読者間の交流が進む場合もあるし、作者の意図しない方向に、読者の興味が盛り上がる場合もある。電子漫画の1つに「コミコ (<http://www.comico.jp/>)」がある。コミコはスマートフォンに最適な漫画作りを進めている。また、読者コメントの積極的な利用も進めている。コミコの運営サイトによると、コメントが作者と読者との相互作用で漫画に影響を与えていると延べている。本研究では、電子漫画に対する読者コメントに着目する。また、特に漫画のキャラクタに着目し、読者がキャラクタに持つ印象を分析する。コメントの収集手法、コメントの統計解析、最後にキャラクタの関係推移について分析手法と結果を述べる。

1. はじめに

現在漫画コンテンツの電子化が進んでいる。従来漫画の表現は、紙面への印刷に適するように進化してきた。近年のスマートフォンおよびタブレットの普及により、電子漫画も増加している。また電子マンガのアプリも数多く提供されている。従来からある紙面のマンガを電子化（多くは画像化）した電子マンガも多いものの、スマートフォンやタブレットでの閲覧に適した表現にのマンガも多い。

ネット上の情報サービスでは、主となる提供コンテンツの提示閲覧だけでなく、視聴者からのコメント投稿と、コメントの提示を行なうサービスが多数存在する。楽天やアマゾン等のEコマースサイトでは、販売商品へのレビューはサービス初期から実装されていた。Eコマースでの商品以外にも、料理紹介やレストラン紹介サイトでは、一般利用者から投稿されたレビューが、その商品やコンテンツへの信頼できる紹介になっている。動画投稿サービスやネット小説サイトでも視聴者コメントは重要である。

我々はニコニコ動画を対象に、動画へのコメントを利用した動画推薦手法の研究[1]と動画再生数の分析[2,3], 楽天での商品レビューの分析[4]について研究してきた。またネット小説サイト「小説家にな

ろう」を対象に、ブックマーク数の分析や、ブックマークのリンク構造を用いた小説推薦を研究してきた。最近では、小説の多様性動向についても分析[5]してきた。

電子コンテンツの場合、利用者（読者、視聴者）と作者との交流が、簡単かつ短時間でやりとり可能になっている。読者のコメントから、読者間の交流が進む場合もあるし、作者の意図しない方向に、読者の興味が盛り上がる場合もある。

本研究では人気の電子漫画サイト「コミコ」(<http://www.comico.jp/>)における読者コメントを分析する。いくつか有るスマートフォンアプリ用電子漫画アプリの中で、コミコアプリは2017年2月現在、日本で2番目に利用者が多い。また、コミコは読者コメントを積極的な利用している。

本研究では電子漫画に対する読者コメントに着目する。漫画への読者コメントを分析することで、漫画作者の意図を読者が理解しているか、その意図を読者が好意的に受け止めているかを知ることができる。読者コメントの漫画作品への影響も分かるかもしれない。本研究では特に漫画のキャラクタに着目し、読者がキャラクタに持つ印象を分析する。コメントの収集手法、コメントの統計解析、最後にキャラクタの関係推移について分析手法を結果を述べる。

2. 対象サイト「コミコ」

先に述べたように、本研究では電子漫画サイトはコミコ (<http://www.comico.jp/>) を研究対象としている。

2.1 コミコの特徴

コミコはNHN comico社が運営する2013年10月設立のサービスである。ニールセン社による2017年2月時点の調査[6]によると、スマートフォン向け電子漫画アプリの中で、コミコは日本国内の利用者数が2位のアプリで、約260万人が利用している(1位はLINE漫画で279万人利用)。

他の電子漫画サービスと異なるコミコの特徴は、過去の漫画を電子化した漫画は提供せずに、コミコ用に作成されたオリジナル漫画の提供を推進すること、スマートフォンでの閲覧に適した漫画を提供している点であろう。漫画は上から下に動かすだけで閲覧でき、紙面の様に左右方向に動かすことは無い。

もう一つの特徴は読者コメントの積極的な利用である。コミコの運営サイトによると、コメントが作者と読者との相互作用で漫画に影響を与えていると述べている。

コミコが提供する漫画は「公式」、「ベストチャレンジ」、「チャレンジ」の3ランクに分けられている。新人の漫画はチャレンジ作品として連載される。人気が出て「ベストチャレンジ」漫画として選出され、そこで品質と人気を認められた作品が「公式」漫画として提供される。表1に各ランクにおける漫画作品の制限を記載する。

表 1: 各ランクの制限

	チャレンジ	ベストチャレンジ	公式
オリジナル	他所で公開した作品も可		オリジナルのみ
1話の最低コマ数	なし	15コマ	30コマ
カラー	モノクロでも良い	カラー	カラー
新話の更新間隔	制限なし	制限なし	隔週

2.2 コンテンツ数

コミコの非公式情報サイト「comico info」[7]に掲載されている2017年8月時点での「公式」漫画作品数を表に示す。

表 2: 漫画作品数

種別	作品数
連載中	260
連載終了	188
(削除済)	9

2.3 タグ

コミコの漫画作品には、利用者が作品検索をしやすくするためにタグが付与されている。タグはジャンルを表すジャンルタグ、公式かチャレンジかを表すランクタグ、それから漫画を表すキーワード的なタグが有る。コミコのタグはTwitterにおけるハッシュタグの様に、#と単語の形をしている。

表3にコミコで用いられている12個のジャンルタグを示す。ほとんどの漫画作品では複数のジャンルタグが付与されている。

表 3: コミコのジャンルタグ (12個)

ドラマ, ギャグ・コメディ, 日常, 学園, 恋愛, ファンタジー・SF, ホラー・ミステリー, アクション, 歴史・時代劇, スポーツ, エッセイ, オムニバス

各ジャンルタグが付与された漫画作品数を表4に示す。複数のジャンルタグを付けた漫画作品が多いため、合計数は表2に示した作品総数より多い。

表 4: 各ジャンルタグの作品数

ジャンルタグ	作品数
ドラマ	167
ギャグ・コメディ	180
日常	154
学園	120
恋愛	136
ファンタジー・SF	140
ホラー・ミステリー	43
アクション	54
歴史・時代劇	10
スポーツ	7
エッセイ	25
オムニバス	3

2.4 コミコ漫画作品の構成

コミコでは、一つの漫画作品に一つのタイトル番号 (title No) が割当てられる。一つの作品は、複数の「話」で構成される。隔週か週間で連載されることを想定しているため、連載中であれば一定期間ごとに新しい「話」が公開される。各「話」は番号(article No)で識別される。

例えば、コミコで最も人気のある漫画作品「ReLIFE」の場合、タイトル番号は2である。タイトル、作者、タグ、話の一覧を示す URL は、

<http://www.comico.jp/detail.nhn?titleNo=2>

である。2017年10月末現在、215話まで公開されている。ReLIFEの215話を示す URL は、

<http://www.comico.jp/detail.nhn?titleNo=2&articleNo=215>

である。

読者は漫画作品に対し、コメントの投稿と応援ポイントが付与できる。コメントは漫画作品の各話に対して投稿される。コメントには投稿者名と投稿日時が付いて、読者に表示される。読者はコメントに対し「いいね」を押して、コメントへの支持をすることもできる。

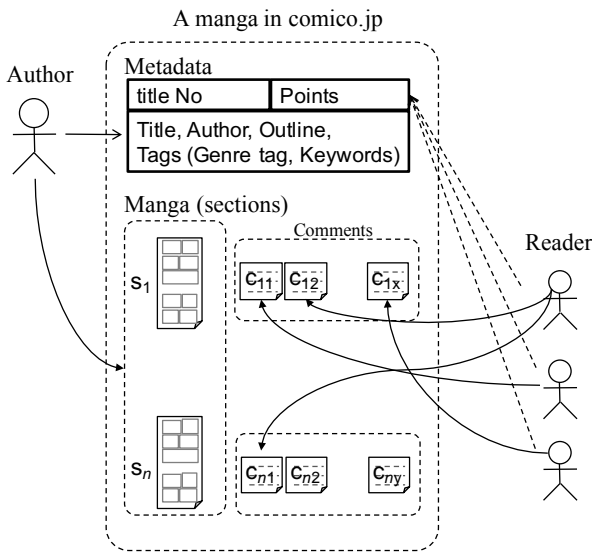


図 1: 漫画, コメント, 応援ポイントの関係

3. コメント分析

3.1 コメント収集

コミコのコメントを機械的に収集するため、Python 言語を用いてコメント収集プログラムを作成した。このプログラムは、漫画作品の番号(title No)とその漫画の最大話数を指定すると、コミコの Web サイトからコメントのページ(HTML)を収集する。入手した HTML ページから、xpath を用いてコメント部分のみをテキストとして出力する。

後述するコメント分析のため、収集コメント群を整形し、検索エンジンに投入した。コメント投稿者名、投稿日時の抽出、コメント本文を形態素解析ツール mecab を用いて単語に分ち書きしたものを作成した。また、コメント本文中に出現する漫画のキャラクター名を抽出する機能も別途作成した。コメ

ント内でキャラクターは略称や愛称で記載されることが多いため、人手で略称や愛称を指定して、キャラクター名を抽出した。

3.2 コメント収集した漫画作品

以下の表に示す5つの漫画作品について、読者のコメントを収集した。

表 5: 分析対象とした5つの漫画作品

titleNo	漫画作品名	作者名	話数
2	ReLIFE	夜宵草	215
20731	血羅の一族	TAMLIN / Cu-rim	81
5803	BEAUTIFUL LEGENDS	redice	91
1308	R!P・G	てっちゃん	163
11585	TWO SOULS ～繋ぎあう心～	tama	111

3.3 キャラクターへの注目

我々は、漫画作品に対する読者の共感が、漫画へ影響すると考えている。読者コメントを分析する際、漫画のキャラクターに着目する。

多くの漫画の原作を書いた小池一夫は、著書[8]の冒頭で「漫画はキャラクターである。キャラクターが起っていれば、漫画はヒットする。」と記述している。また「魅力的なキャラクターを創ることで、キャラクターは読者に愛されます。『あのキャラに会いたい』と思ってもらえると、作品が売れ、続きが出版されます。」とも述べている。この意見に我々も同意し、読者が漫画のキャラクターに共感しているかを調査する。

本研究では、読者コメント内にキャラクター名が出現するならば、コメント者はそのキャラクターに何らかの共感があるとする。人が何かに共感して、それを文章で表現する場合、自分が好む、あるいは応援したい人・モノの名前を記述するだろう。文章ではない実世界のスポーツ・芸能・政治イベントを見ると、応援対象の名前を連呼したり、プラカードで示すことも多い。つまり、ある対象の名前を記述することは、記述者がその名前を持つ実体へ興味や関心が有ることを示している。

3.4 キャラクターの出現頻度

あるキャラクターへの読者の共感を定量的に測るため、コメントでの出現回数を考える。コミコでは漫画作品の「話」毎に読者がコメントを投稿できる。第 i 話のコメント群を $C_i = \langle c_1, c_2, \dots \rangle$ とする。た

だし、 C_i に含まれるコメントは漫画の「話」が公開されてから1週間以内に投稿されたコメントだけを分析対象とする。コミコでは一つの「話」の公開日からずいぶん後でも、その話へのコメントを投稿可能である。本研究では連載におけるコメント動向に注目しているため、分析対象のコメントも漫画作品の「話」が公開されてから、1週間以内のコメントだけを分析対象とした。

キャラクタ K の出現回数に3つ考えた。

- (1) 全コメントでの K の出現数: $\sum \text{tf}(K, c_j)$
- (2) K が出現したコメント数: $\text{df}(K)$
- (3) キャラクタ K が出現したコメントを記述したユーザ数

ほとんどのコメントは短いため、キャラクタ名を連続記述するコメントは少ない。また1人の読者が、1つの「話」に複数コメントを投稿することは少ないため、上記(2)を、一つの「話」におけるキャラクタ K の出現数とした。

また、キャラクタペアの出現頻度も着目する。漫画のストーリー展開で、2人のキャラクタが親友になることや、恋人になる場合がある。スポーツの対戦相手の関係もある。恋愛を扱う漫画では、2人の関係進展を読者が興味を持ち共感することが大事であろう。読者が共感していれば、その漫画は人気作品になるだろう。

2人の関係への読者の共感を定量的に測るため、キャラクタ K_1, K_2 の共起頻度を調べる。

- (4) K_1 と K_2 が共起するコメント数: $\text{df}(K_1, K_2)$

4. ReLIFE コメントの分析

コミコで最も人気が高く、そのため読者も多い漫画作品「ReLIFE」の読者コメントについての分析結果を述べる。ReLIFEは「受験や就職につまづき続け無職となった主人公が社会復帰の実験として高校生活をやり直す姿をコミカルに描いた作品[9]」である。

表 6: ReLIFE の主要キャラクタ

No	氏名	読み	性別
1	海崎 新太	かいざき あらた	男
2	日代 千鶴	ひしろ ちずる	女
3	夜明 了	よあけ りょう	男
4	狩生 玲奈	かりう れな	女
5	大神 和臣	おおが かずおみ	男
6	小野屋 杏	おのや あん	女

表 6 に ReLIFE の主要キャラクタ名を表に示す。高校3年生なので見た目の年齢は全員17歳である。ただしキャラクタ1と2は大人が薬で若返っているという設定である。

4.1 コメント数の推移

図2に、ReLIFEの各話に対し投稿されたコメント数を示す。図2では目立つ部分にA~Fを付けている。初期は人気が出ていないためコメント数は少ない。図2のAの25~27話の公開時期には、アプリ内でコメントを用いたイベントが開催されていた。Bの39と40話では作者が書籍化通知をしており、それへの反応で増加している。Cの118話ではコメント数が最高になる。Dの144話もコメントが多い。CとDは物語が盛上がる部分であるためコメント数に反映している。Fの前後になると、一週間のコメント数は減少する。累積コメント数は少なくない。Fの187話では映画化が通知され、それへの読者反応がコメント数に出ている。

作者の意図するストーリーや、サイト運営が仕掛けたイベントに読者が反応してコメント数が増えることが分かる。

4.2 キャラクタの出現頻度

人手で ReLIFE の各キャラクタについて別名や愛称をコメント群から抜き出した。それを用いて各キャラクタの $\text{df}(K)$ を数えた。表6の主要キャラクタ No.1~6 の、話毎の頻度を図3のグラフに示す。グラフの青線は文書頻度で、橙線は頻度をコメント数で割った値で、キャラの出現コメントの割合になる。

主人公の1.海崎と主ヒロインの日代は全体的に頻度が高い。第27話でサブキャラの5.大神が急増している。漫画内の状況が人気に繋がり、かつイベント時の悪ノリで増加したのだろう。

4.3 キャラクタペアの共起頻度

ReLIFEのコメントから2人のキャラクタ K_1 と K_2 が共起頻度 $\text{df}(K_1, K_2)$ を調査した。No1~6のキャラクタのペアであるため全15ペアの共起頻度を調査した。調査の結果15ペアのうち、共起頻度が高い表7に示す8ペアのグラフを示す。

表 7: 共起頻度の高いキャラクタペア

No	ペア	性別
1-2	海崎と日代	男-女
1-3	海崎と夜明	男-男
1-4	海崎と狩生	男-女
1-5	海崎と大神	男-男
2-3	日代と夜明	女-男
2-4	日代と狩生	女-女
3-6	夜明と小野屋	男-女
4-5	狩生と大神	女-男

主人公とヒロインの関係である 1-2 (海崎-日代) は共起頻度が高い。110 話頃から 2 人の恋愛が描かれると頻度が上がる。また 4-5 (狩生-大神) は、も

96~113 話で 2 人の恋愛が描かれている。その期間の共起頻度が高い。読者は物語内のカップル関係に大きな共感を示すと言える。

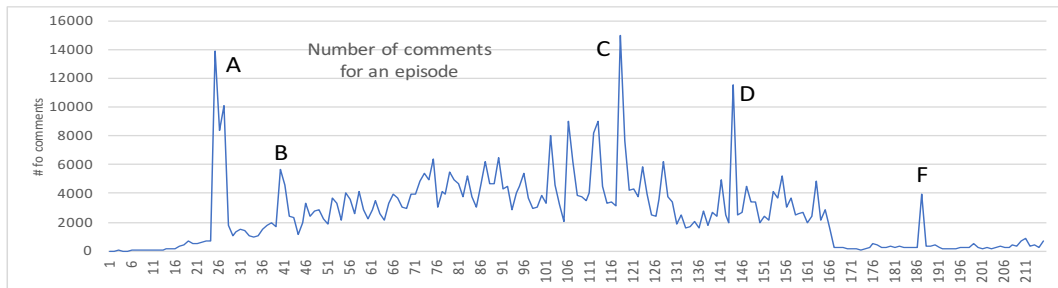


図 2：各話への投稿コメント数 (1 話：2013.10.17-215 話：2017.10.20)

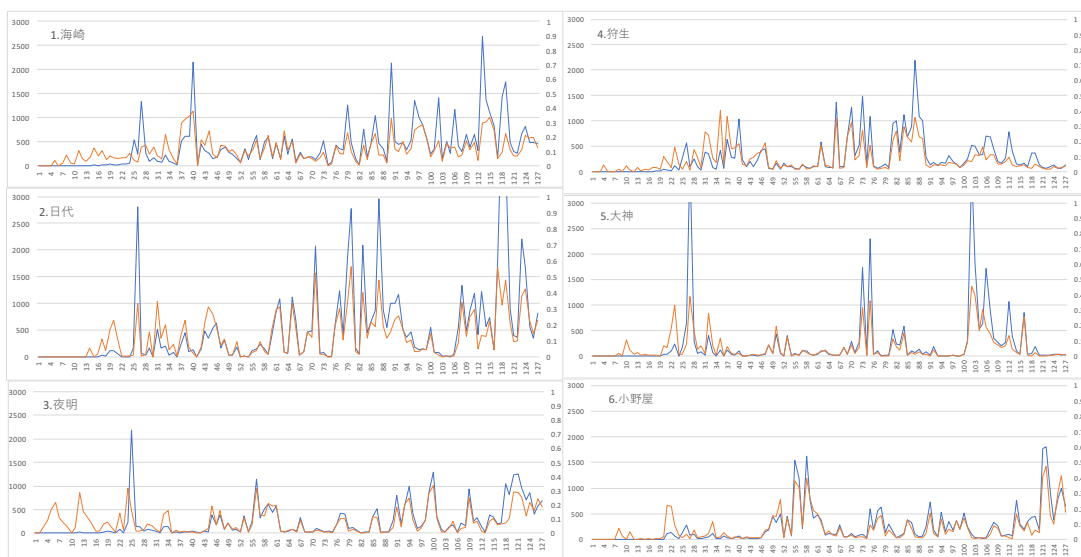


図 3：各話コメントでのキャラクターの頻度 (1 話：2013.10.17-127 話：2016.03.04)

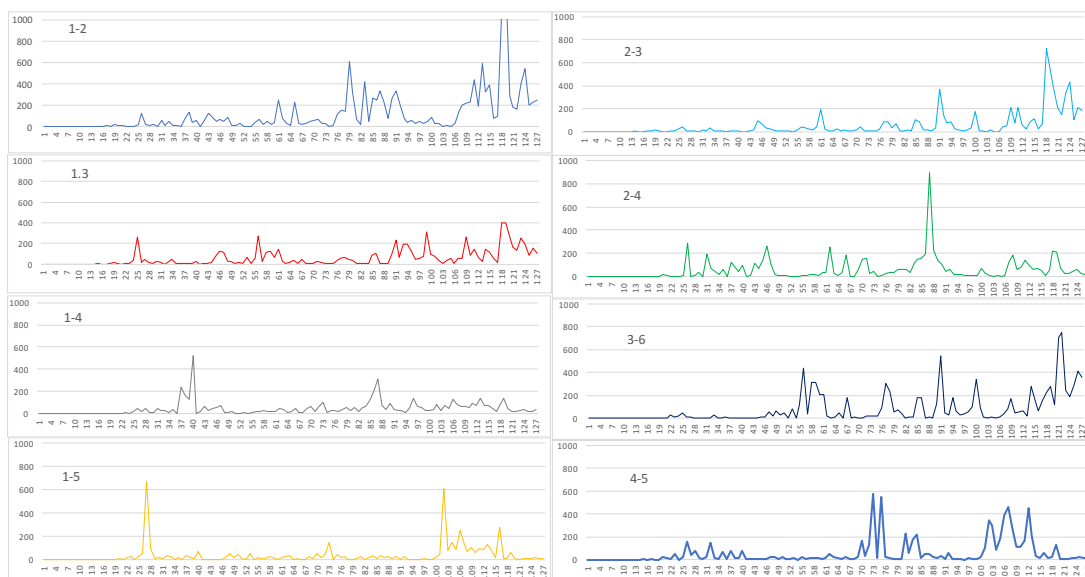


図 4：各話コメントでのキャラクターの共起頻度 (1 話：2013.10.17-127 話：2016.03.04)

読者コメントにおける、キャラクターの頻度や、ペアの共起頻度を観測することで、そのキャラクターおペアへの共感（反感も含む）の動向を計測できる。ある閾値より頻度が上昇・下降した際の通知や、頻度の急増・急減の際の通知を行えば、漫画作品の作成や編集の支援になる可能性がある。急増や急減を検出するには、前話との頻度比を算出すれば良い。

5. 属性選択によるコメント分析

次に属性選択手法[11]によるコメント分析について述べる。今回は具体的な結果を得ていないため、目的と手法を述べる。

5.1 属性選択を用いる理由

4節で述べたように、読者コメントから、キャラクターへの共感（反感も含む）の動向を計測できる。4節で述べた ReLIFE は漫画もコメントも詳細に読んだので、読者が何に共感したのかを理解した。しかし全ての漫画作品を人手で分析するのは効率的ではない。読者が共感する理由を機械的に抽出できれば、漫画の作者や編集者の支援になる。そこで、あるキャラクターのどこに読者が共感しているのかを、属性選択手法で解析する。

廣川は、機械学習手法 SVM (Support Vector Machine) における属性選択について研究してきた[10]。SVM は多量の属性からなるデータを用いた分類で優れた性能を示すと言われている。一方、属性を増加しすぎると推定性能が悪化すると示す研究や、多くの属性を用いた場合と同程度の推定性能を少数の属性でも実現できると示す研究も有る。廣川らの分類問題における属性選択の目標は、多属性を用いた場合と同程度の推定性能を少数属性で実現することで、少数属性の方が多数属性を用いるよりも推定性能が高くなることではない。属性選択を適切に行うことで全属性を用いた場合と同程度の推定性能を少数属性でも実現している[10]。

5.2 属性選択による共感属性抽出

ある第 i 話のコメント群 C_i でキャラクター K がコメントに頻出したとする。各コメント $c \in C_i$ について、コメント者名、投稿日時、コメント内のキャラクター名と単語の頻度のように、複数属性の数値から成るベクトルで表現する。コメント群 C_i を、 K を含むものと含まないものに分け、正例を K を含むコメント、負例を含まないコメントとして SVM による分類器を作成する。

属性選択手法を用いて、重要度が大きい属性を見つける。漫画作品へのコメント群を、読者がキャラクターあるいは漫画作へ共感したものと、そうでない

ものに SVM で分類する。SVM 適用の際、なるべく多くの属性を用いる。その後、属性選択の手法を用いて、重要属性を抽出する。

5.3 前話コメントの次話の影響

前回の「話」の影響も考慮するため、前回の話に対するデータを、今回の話の属性として用いる。連載漫画作品の場合、第 $i+1$ 話は、前回の第 i 話の続きであるため、ストーリーの関係がコメントに出る可能性がある。第 i 話の読者コメントに作者が影響を受け、コメント内容を第 $i+1$ 話に反映する可能性もあるため、その状況を読者コメントから抽出したい。

5.4 人手による共感度

可能であれば、各コメントが、どの程度共感を示した文章であるかを人手でスコア付けしても良い。以下の意味で -1 から 1 の範囲で共感度スコアを与えると、反感コメントの抽出にも役立つ。

- 1 : その「話」にコメント者が完全に共感
- 0 : その「話」への共感無し
- -1 : その「話」にコメント者が完全に反感

6. 関連研究

Murakami らは文献[11]で、漫画の紙面（画面）から、漫画内の人物キャラクターと文章を人手で入力し、と、出現頻度分析でその人物の性格を表す文字を結び付ける手法を提案している。また、共起頻度から人物間の関係を相関図に表す手法も提案している。我々も漫画のキャラクターに着目した頻度解析を行っているものの、解析対象が漫画ではなく、読者コメントである所が異なる。

Cohn らは文献[12]で、日本の漫画に特徴敵に出現する「Visual vocabulary (視覚的語彙)」を分析している。従来から日本漫画で用いられてきた 73 個の視覚的表現を抽出している。これらが、少年漫画・少女漫画から選んだ 20 冊でも用いられて居ることを示している。我々の研究対象である読者コメントには多数の絵文字が含まれているものの、本論文では絵文字は分析していない。絵文字を分析する場合、Cohn らの研究を援用できる。

Dharma らは文献[13]で SNS における通信に漫画を用いる方法について研究している。SNS は世界最大のコミュニティであり、多種の双方向通信や多様なニュース配付を実現している。Dharma らの提案する漫画を用いた通信は、SNS 利用者に趣味分野の満足度を向上すると述べている。本論文が対象とするコミコの読者コメントは、読者の思いを一方向的に述

べるものであるため、読者間、あるいは作者と読者の間のやり取りには用いられない。

7. おわりに

本研究では人気の電子漫画サイト「コミコ」の読者コメントを分析した。コメントを機械的に収集するプログラムを Python 言語を用いて作成した。コメント分析の際、漫画の登場人物であるキャラクタに特に着目した。収集したコメントから、キャラクタの文書頻度を数えた。漫画を含むストーリーでは人間関係が重要である。そこで2人のキャラクタの共起頻度も数えた。

その結果、読者は「主人公・ヒロイン」や「恋心」という観点でキャラクタへの共感を持ちやすく、さらにキャラクタペアでは「カップル」という関係に一番共感を持つことがわかった。またキャラクタペアにおいてはただの「友人」、「仲間」ではなく物語の中で「敵対」から「親友」へと変化した関係性に共感を持つこともわかった。

今後は分析範囲を多くの漫画へ拡大する予定である。人気の漫画作品の全コメントを収集して分析することで、読者と漫画キャラクタの関係を抽出したい。また、本論文では方法のみを述べた属性選択手法による分析も行いたい。これらの分析を終えた後、読者の共感が連載漫画の作者に与える影響を定量的に示す事を確認したいと考えている。

謝辞

本研究は JSPS 科研費 15K00451 の助成による。

参考文献

- [1] Naomichi Murakami, Eisuke Ito: Emotional video ranking based on user comments, Proc. of ACM iiWAS2011, pp.499-502, (2011)
- [2] Kazuhisa Noguchi, Tomoya Iida, Eisuke Ito: An analysis of CGM contents pageview using SIR Model and GBM,

Proc. of ICCTD2017, pp.19-21 (2017)

- [3] Kyohei Kamihata, Eisuke Ito: A quantitative contents diversity analysis on a consumer generated media site, Proc. of AROB 21st 2016, pp.436-440, (2016)
- [4] 林驍, 伊東栄典, 廣川佐千男: 大規模 EC サイトの商品レビュー傾向分析, 情報処理学会 情処研報 2015-ICS-181, No.7, pp.1-6, (2015)
- [5] 本田優也, 伊東栄典: 利用者投稿型小説サイトにおけるキーワードの多様性分析, 情報処理学会 火の国情報シンポジウム 2017, A2-4, (2017)
- [6] ニールセン社: 若年層でスマートフォンからマンガを読む習慣が定着, http://www.netratings.co.jp/news_release/2017/03/Newsrelease20170328.html, (2017) (Accessed at Oct.30, 2017)
- [7] Comico info: <https://comico.info/comic/> (Accessed at Oct.30, 2017)
- [8] 小池一夫: 小池一夫のキャラクター新論 ソーシャルメディアが動かすキャラクターの力, ゴマブックス (2017)
- [9] Wikipedia “ReLIFE”: <https://ja.wikipedia.org/wiki/ReLIFE> (accessed at Nov.02, 2017)
- [10] T. Sakai and Sachio Hirokawa: Feature words that classify problem sentence in scientific article, Proc. of the 14th International Conference on Information Integration and Web-Based Applications and Services, pp. 360-367 (2012).
- [11] Harumi Murakami, Ryota Kyogoku, Hiroshi Ueda: Creating character connections from manga, Proc. of ICAART 2011, pp. 677-680 (2011).
- [12] Neil Cohn, Sean Ehly: The vocabulary of manga: Visual morphology in dialects of Japanese Visual Language, Journal of Pragmatics, Vol.92, pp.17-29 (2016)
- [13] Anak A. G. Dharma, Hiromitsu Kumamoto, Shogo Kochi, Natsuki Kudo, Wei Guowei, Chiu Shu-Chuan, Kiyoshi Tomimatsu: The utilization of Social Networking Service and Japanese manga in strategic user generated design, Proc. of ICEEI 2011, pp.1-6 (2011)