

科学研究費助成事業 研究成果報告書

平成 29 年 8 月 4 日現在

機関番号：54501

研究種目：若手研究(B)

研究期間：2015～2016

課題番号：15K21592

研究課題名(和文)感情推定アルゴリズムに基づくオンラインでの性格推定に関する研究

研究課題名(英文)A Personality Estimation in Customer Generated Media based on Emotion Judgment

研究代表者

奥村 紀之(Okumura, Noriyuki)

明石工業高等専門学校・その他部局等・講師

研究者番号：40510277

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：本研究では、SNSなどに投稿された記事の内容から、著者の見せかけの性格を推定するための手法について検討した。特に、Blogのように推敲を重ねて公開される記事と、Twitterのようにその場の思いつきで投稿される記事とでは、表面的に現れる個性、性格に大きな違いがあることを被験者実験により確認している。また、見せかけの性格と著者本来の性格の間にも大きな差が生じていることを確認している。本手法を自動化するために感情推定アルゴリズムとの相関分析を行い、Big Fiveの5因子との関連を抽出している。

研究成果の概要(英文)：In this research, we examined the method to estimate a virtual personality in Customer Generated Media based on the articles submitted to SNS. We especially confirmed the difference between blogs submitted after arranging author's verbs and objects, and tweets submitted without arranging. We also found out the difference between author's virtual personality and author's real personality. Finally, we extracted the relationship between the five factor in Big Five Model and emotion judgment algorithm to automatize personality estimation method.

研究分野：知識工学

キーワード：性格推定 感情推定 顔文字

1. 研究開始当初の背景

オンラインでのテキストベースのコミュニケーションが浸透してきている一方で、不要な発言によるトラブルが多発している問題があった。特に、自分の投稿の読み手に対してどのような印象を与えているかという点を適切にユーザが把握できていないという問題があった。投稿者は可能な限り意識的に自分の発言(投稿)に気を配っているが、必ずしも読み手が意図通りに解釈するとは限らない。こういった問題を解決する必要があった。

2. 研究の目的

本研究の目的は、ソーシャルメディアにおける、ネットワーク上での、自己の、他者から見た性格を定量評価し、自分自身の本来の性格との差分を抽出することで、コミュニケーションの円滑化を図るというものである。

3. 研究の方法

本研究の実施方法は大きく2つあり、1つ目は、感情解析のための顔文字の分析、2つ目は、既存の感情推定アルゴリズムを利用した性格推定のための調査である。顔文字については、使用状況、文脈など種々の要因で含有する感情情報に変化するという問題に対処するため、大規模な顔文字の辞書を構築し、分析を行った。性格推定に関しては、Macromillを利用したアンケート調査により、Twitter、Blogの双方を利用しているユーザに対して、Big Fiveに基づく性格診断を実施し、ユーザ本来の性格、ならびにTwitter、Blogの投稿を収集し、分析を行った。

4. 研究成果

本研究の成果は大きく2つある。1つ目は、投稿者本来の性格と投稿された記事から推定される見かけの性格には差分があることを明らかにしたこと、2つ目は、大規模な顔文字辞書の構築により、顔文字には変化の基本となる原形が存在することを確認したことである。

まず1つ目の成果に関しては、雑誌論文⑥ならびに、学会発表⑦で報告している。この研究で明らかにしたことは、TwitterやBlogの記事から直接投稿者本来の性格を推定するというタスクに対して、TwitterやBlogへ投稿された記事に基づいて形成される投稿者の見かけの性格について検証をしている。本実験では、被験者実験により、BlogやTwitterの投稿を熟読し、NEO-FFIの性格検査を、そのブログの投稿者であればこのように回答するはずだという視点から回答させ、投稿者の擬似的な性格を抽出している。

被験者実験では、5名の被験者に記事を読ませ回答させている。図1には、被験者にTwitterへの投稿記事を読ませた際の性格推定の結果を、図2にはBlogを読ませた際の性格推定の結果をそれぞれ示している。

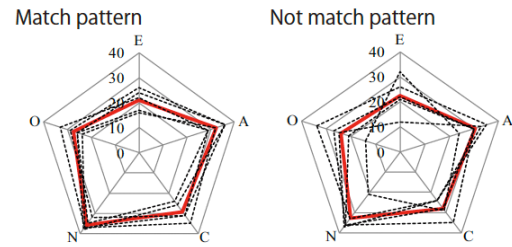


図1: Twitterからの見かけの性格推定

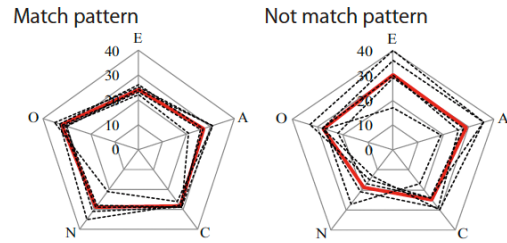


図2: Blogからの見かけの性格推定

図1、図2に示すように、5名の被験者がほぼ同一の傾向を示す回答をしたパターンと、5名の被験者の回答が一致しなかったパターンが得られた。23名分のTwitterとBlogに対して、Twitterでは16名分の記事について5名の被験者が同一の性格を推定し、Blogでは12名分の記事について同一の性格を推定するという結果になった。

この要因として、Twitterに投稿される記事には比較的感情的な成分が含まれており、投稿者の気持ちなどを汲み取りやすかったのではないかと考えられる。一方、Blogの記事は投稿前に推敲してから投稿することが多いため、感情的な表現や投稿者の主義主張が見え隠れするような記述が抑制されていると考えられるため、5名の被験者が一致するパターンが少なかったものと考えられる。このことは、表1、表2に示す感情成分と5因子との相関関係により類推される。

表1: Twitterの投稿に含まれている感情成分とBig Fiveの各因子との相関

	E	A	C	N	O
Anger	-0.03	-0.17	0.62	0.15	-0.13
Sadness	0.08	-0.10	0.22	-0.06	-0.21
Fear	-0.13	-0.17	0.51	0.08	-0.04
Disappointment	-0.15	0.51	0.67	0.65	0.20
Regret	0.14	0.29	0.49	0.25	0.32
Guilt	0.04	-0.33	0.31	-0.12	-0.23
Shame	0.16	-0.17	0.33	0.04	-0.06
Pleasure	0.24	0.27	0.35	0.20	0.06
Ease	0.10	0.26	0.79	0.43	-0.06

表1、表2から、Twitterの方が感情に対して強い相関を示していることが分かる。このため、感情推定アルゴリズムを利用した見かけの性格推定には、Twitterの投稿の方が適しているということが分かる。一方で、Blogの記事から見かけの性格を推定する場合には、別途観点を抽出していく必要がある。

表 2 : Blog の投稿に含まれている感情成分と Big Five の各因子との相関

	E	A	C	N	O
Anger	-0.18	-0.29	-0.25	0.10	-0.36
Sadness	0.01	0.12	0.16	0.03	-0.16
Fear	-0.27	-0.30	-0.22	0.17	-0.31
Disappointment	0.16	0.14	0.03	-0.20	-0.41
Regret	-0.03	0.08	-0.18	0.13	-0.08
Guilt	-0.34	-0.31	-0.20	0.29	-0.14
Shame	0.14	0.19	-0.02	-0.11	-0.21
Pleasure	0.10	0.21	-0.02	-0.04	-0.18
Ease	-0.07	-0.02	-0.15	0.18	-0.24

2 つ目の成果として、顔文字に関する大規模な分析結果が挙げられる。当初、感情推定アルゴリズムを拡張するために顔文字を導入する目的で研究を開始したが、顔文字には道の情報が大量に含まれている可能性があることから、顔文字そのものの分析を行っている。

顔文字に関しては、第 30 回人工知能学会全国大会、および第 31 回人工知能学会全国大会においてオーガナイズドセッションを企画し、多数の発表と討論を行ってきている。その結果を解説論文として雑誌論文①～④にまとめ、掲載されている。

顔文字の大規模な辞書では、顔文字自体は約 7 万種類登録されており、そのうち 4 万種の顔文字についてアノテーションを実施している。これにより、例えば、「#」という記号が付与された顔文字はほぼ例外なく「怒り」の感情を持つことや、「*」が付与された顔文字は、元の顔文字の感情成分を強調する効果があることなどが分かってきている。

また、このような顔文字の変位を顔文字の活用形として考え、顔文字の中心をなす顔文字の原形を抽出することに成功している。顔文字の原形は現在のところ 3110 種存在しており、分類方法によっては 3110 種をさらに削減することも可能である。

これらの顔文字の原形が適切に推定することができれば、未知の顔文字であっても感情成分を抽出するなど分析が容易になることに加え、顔文字の自動生成といったタスクにも応用可能である。これらの大規模辞書に関しては、現在公開に向けて調整しているところである。

5. 主な発表論文等 (研究代表者は下線)

[雑誌論文] (計 6 件)

- ① 奥村 紀之・プタシンスキ ミハウ・ジェプカ ラファウ、工学的・心理学的側面から見た顔文字の分析に関する研究動向、人工知能学会誌、解説論文、Vol. 32, No. 3, pp. 342-349, 2017
- ② プタシンスキ ミハウ・奥村 紀之・ジェプカ ラファウ、顔文字の現象および研究の概観-記号の遊びが科学されるよ

うになった道-, 人工知能学会誌, 解説論文, Vol. 32, No. 3, pp. 333-341, 2017

- ③ ジェプカ ラファウ・奥村 紀之・プタシンスキ ミハウ、複数の世界をつなぐ顔-現代における絵記号の意味と可能性-, 人工知能学会誌, 解説論文, Vol. 32, No. 3, pp. 350-355, 2017
- ④ 藤澤 日明・松本 和幸・奥村 紀之・吉田 稔・北 研二、アスキーアートへの挑戦-画像特徴量によるアプローチ-, 人工知能学会誌, 解説論文, Vol. 32, No. 3, pp. 364-370, 2017
- ⑤ Akihiro Toyoshima・Noriyuki Okumura, A Comparison of Concept-base Model and Word Distributed Model as Word Association System, 20th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems KES2016, Procedia Computer Science, 査読あり, Volume 96, pp. 385-394, 2016
- ⑥ Noriyuki Okumura・Manabu Okumura, A Construction of Knowledge Base for Personality Estimation based on Submitted Text Data in Twitter or Blogs, In Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - Volume 2: KEOD, 査読あり, pp. 418-423, 2015

[学会発表] (計 9 件)

- ① 大町 凌弥・奥村 紀之、文章と顔文字の組み合わせによる感情推定、第 31 回人工知能学会全国大会、202-0S-22a-2, 2017
- ② 奥村 嶺・奥村 紀之、ニューラルネットワークによる顔文字の原形推定、第 31 回人工知能学会全国大会、203-0S-22b-1, 2017
- ③ 奥村 紀之、分類器による顔文字の原形推定、信学技報, vol. 116, no. 379, NLC2016-37, pp. 93-96, 2016
- ④ 奥村 紀之、言語解析のための大規模顔文字辞書、第 30 回人工知能学会全国大会、3H3-0S-17a-1, 2016
- ⑤ 奥村 紀之、感情抽出のための顔文字の原形推定、信学技報, vol. 116, no. 78, NLC2016-1, pp. 1-4, 2016
- ⑥ 奥村 紀之、顔文字の原形抽出、言語処理学会第 22 回年次大会、P1-1, 2016
- ⑦ 奥村 紀之・奥村 学、オンラインでの振る舞いから想定される人物像の特徴、信学技報, vol. 115, no. 70, NLC2015-13, pp. 73-77, 2015
- ⑧ 香川 夏美・奥村 紀之、被験者実験による常識判断知識ベースの整合と検証、信学技報, vol. 115, no. 70, NLC2015-3, pp. 13-18, 2015

- ⑨ 豊嶋 章宏・奥村 紀之, 語の連想体系としての概念ベースの評価, +, 2015

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
取得年月日：
国内外の別：

[その他]

ホームページ等

<http://broccoloid.jp/>

6. 研究組織

研究代表者

奥村 紀之 (OKUMURA, Noriyuki)

明石工業高等専門学校・電気情報工学

科・講師

研究者番号： 40510277