

Q&A サイトにおける自分の投稿に対する評価を操作するために

複数のアカウントを利用するユーザの検出

Detection of Multiple Account Users Suspected of

Manipulating Evaluations of Their Submissions in a Community Site

石川 尚季 (いしかわ なおき・Naoki Ishikawa)¹・梅本 顕嗣 (うめもと けんじ・Kenji Umemoto)²・
西村 涼 (にしむら りょう・Ryo Nishimura)³・渡辺 靖彦 (わたなべ やすひこ・Yasuhiko Watanabe)⁴・

岡田 至弘 (おかだ よしひろ・Yoshihiro Okada)⁵

^{1 3} 龍谷大学大学院 理工学研究科 情報メディア学専攻 修士課程・

² 龍谷大学 理工学部 情報メディア学科 学生・

⁴ 龍谷大学 理工学部 情報メディア学科 講師・⁵ 龍谷大学 理工学部 情報メディア学科 教授

[Abstract]

Some users in a community site abuse the anonymity and attempt to manipulate communications in a community site. These users and their submissions discourage other users, keep them from retrieving good communication records, and decrease the credibility of the communication site. To solve this problem, we conducted an experimental study to detect users suspected of using multiple user accounts and manipulating evaluations in a community site. In this study, we used messages in the data of Yahoo! chiebukuro for data training and examination.

[キーワード]

多重アカウント, コミュニティサイト, なりすまし, 心豊かなコミュニケーション

1. はじめに

Google などの検索エンジンによって、インターネット中の情報に誰にでもアクセスできる環境が整いつつある。しかし、検索用語を1つ1つ考えて検索エンジンで調べていく検索方法は、誰にとっても簡単なわけではない。調べる対象について十分な知識をもっていない人や、情報の取り扱いになれていない人にとっては、Yahoo! 知恵袋などのQ&A サイトで質問する方が簡単で、しかもよい内容の回答を得られることも多い。

検索エンジンを利用するより簡単に情報を Q&A サイトで収集できるのは、質問に対して回答を積極的に投稿しようとするユーザがいるからである。したがって、そうしたユーザがなぜ回答を投稿しようとするのかを検討することは重要である。三浦らは、Yahoo! 知恵袋で回答者が投稿する理由の第一が質問者を助けたいという利他的理由であることを統計調査によって明らかにした[1]。見ず知らずの他人を繰り返し助けようとする回答者が多いのは、Q&A サイトで「助けてほしい」と考えている人と「だれかのためになれたら」と考えている人が出会い、心豊かなコミュニケーションが行われ、その経験がユーザたちを励まし、新たなコミュニケーションに関わることを促しているのではないかとわれわれは考えている。

こうした心豊かなコミュニケーションでは匿名であることが重要である。匿名のコミュニケーションであることで、照れ、恥かしさ、評判などを気にしないで質問、相談、回答、議論などをすることができるからである。しかし、匿名性を悪用して、Q&A サイトにおけるコミュニケーションを妨害しようとするユーザもいる。したがって、Q&A サイトにおける心豊かなコミュニケーションを促進するためには匿名性を悪用するユーザを検出することが重要である。

匿名性の悪用の1つに、複数のアカウントを利用して、1人のユーザによる投稿を複数のユーザからの投稿に

¹ t10m096@mail.ryukoku.ac.jp

² t070400@mail.ryukoku.ac.jp

³ r_nishimura@afc.ryukoku.ac.jp

⁴ watanabe@rins.ryukoku.ac.jp

⁵ okada@rins.ryukoku.ac.jp

表-1 Yahoo!知恵袋の「パソコン、周辺機器」、「病気、症状、ヘルスケア」、「政治、社会問題」、および全カテゴリ (286 カテゴリ)における質問・回答のメッセージの投稿者の異なりと投稿件数

カテゴリ	質問者数	質問件数	回答者数	回答件数
パソコン、周辺機器	43493	171848	27420	474687
病気、症状、ヘルスケア	29954	84364	38223	289578
政治、社会問題	13259	78777	25766	403306
全カテゴリ (286 カテゴリ)	165064	3116009	183242	13477785

見せかけようとするものがある。他人のふりをする著者を機械学習および統計的手法によって識別する方法はこれまで数多く研究されてきた[2][3][4][5][6][7]。しかし、Q&A サイトへの投稿の評価を操作しようとして複数のアカウントを利用するユーザを検出する方法はまだあまり研究されていない。そこで本研究では、複数のアカウントを利用して質問とその回答を繰り返し Q&A サイトへ投稿し、回答者としての評価を操作しようとするユーザを、投稿履歴と文体の特徴を手がかりに異常な投稿行動の検出と文体の類似判定によって検出する方法を提案する。

複数のアカウントを利用して不適切な投稿を行うユーザについての情報(人数、目的、方法など)は公開されていない。このため、提案手法の精度を議論することはむずかしい。そこで本研究では実験結果を詳しく示し、提案手法の有効性について議論する。本研究で検出された疑わしいユーザについての情報は、複数のアカウントを利用してコミュニティサイトでのコミュニケーションを操作しようとするユーザの目的やふるまいを検討する手がかりになると考えられる。

2. 自分の投稿への評価を操作するために複数のアカウントを利用するユーザ

本研究では、Q&A サイトの例として Yahoo!知恵袋を取り上げる。Yahoo!知恵袋は、登録されたユーザが質問と回答のメッセージをやりとりすることができる Q&A サイトである。この Yahoo!知恵袋に 2004 年 4 月から 2005 年 10 月までに投稿された質問約 311 万件、回答約 1347 万件のメッセージが国立情報学研究所から公開されている⁶。この公開データに収録されている質問には 1 件以上の回答が必ずあり⁸、そのうち 1 件がベストアンサーに選ばれている。この Yahoo!知恵袋のデータの「パソコン、周辺機器」「病気、症状、ヘルスケア」「政治、社会問題」および全カテゴリ (286 カテゴリ)に質問あるいは回答を投稿したユーザの異なりと投稿件数を表-1 に示す。

Yahoo!知恵袋では、質問と回答を匿名で投稿することができる。匿名での投稿は、照れ、恥かしさ、評判などを気にしないで質問、相談、回答、議論などを行うことができるので、心豊かなコミュニケーションを促進するには重要である。しかし、匿名で投稿した質問あるいは回答であってもアカウントによってリンク可能であるので、投稿を重ねれば個人を識別できる情報が蓄積されるおそれがある[8]。このため、匿名性を悪用するためではなく、個人を識別されないために複数のアカウントを利用しているユーザも少なくないと考えられる。一方、自分の投稿は優れた内容であると思われたい、そして優れた投稿をする自分は優れた投稿者であると思われたいために複数のアカウントを利用するユーザもいると考えられる。そうしたユーザは、具体的には、

- 他人のふりをして、自分の投稿により評価を与える
- 他人のふりをして、自分の投稿を支持する投稿をする
- 複数のユーザが反対しているふりをして、他人の投稿に反対する投稿をする

ことが考えられる。このような投稿を繰り返すユーザがいると、まじめに議論しようとするユーザは新たにメッセージを投稿しようとする意欲を失い、結果としてその Q&A サイトは利用されなくなるおそれがある。また、他人のふりをして自分の投稿により評価を与えるユーザは、Q&A サイトに投稿された質問と回答を知識として 2 次利用する場合にも問題になる。投稿内容の評価が操作されている Q&A サイトの質問と回答の内容は信頼できない

⁶ <http://chiebukuro.yahoo.co.jp>

⁷ <http://research.nii.ac.jp/tdc/chiebukuro.html>

⁸ 回答が投稿されなかった質問は削除されているので、この公開データに収録されている質問には 1 つ以上の回答が必ずある。

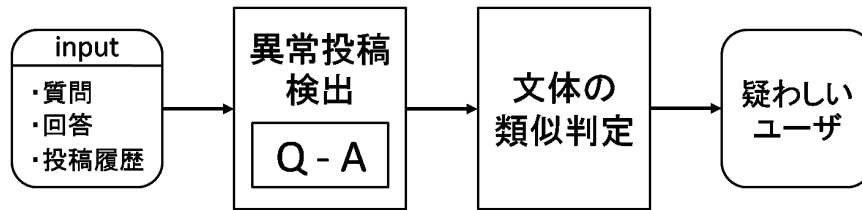


図-1 複数のアカウントを用いて回答者としての自分の評価を操作しようとするユーザを検出する方法

からである。したがって、自分は優れた投稿者であるとみせかけようとして複数のアカウントを利用するユーザ、特に他人のふりをして自分の投稿により評価を与えるユーザを検出することは重要である。そこで本研究では、他人のふりをして自分の投稿により評価を与えることを繰り返し行うユーザを検出する方法について検討する。

Yahoo!知恵袋では、優れた回答を評価する仕組みとしてベストアンサーの判定がある。ベストアンサーとは各質問の投稿者(質問者)によって最もよいと判定されて選ばれた回答である。このため、投稿した回答がベストアンサーに選ばれる割合(ベストアンサー率)が高い投稿者は優れた投稿者(回答者)であると考えられることが多い。そこで、不適切な方法を用いても自分が優れた投稿者(回答者)であると思われたいユーザは、他人のふりをして投稿した質問に自分で回答し、その回答をベストアンサーに選ぶことを繰り返し行うことが考えられる。こうしたユーザを検出するために、本研究では以下の3つの予想をたてた。

(予想 1) できるだけ少ない投稿回数で自分の評価を操作しようとする

(予想 1) が正しければ、1 人のユーザが 2 つのアカウント、ユーザ i とユーザ j を用いてできるだけ少ない投稿回数で回答者役のユーザ(ユーザ j とする)の評価を高めようとする、

1. ユーザ j はユーザ i の質問に対して異常に多くの回答を投稿する
 2. ユーザ i はユーザ j から異常に多くの回答を受け取る
 3. ユーザ i はユーザ j の回答に対して異常に多くのベストアンサーを与える
- の異常な投稿が行われることが考えられる。

(予想 2) 複数のカテゴリで自分の評価を操作しようとする

自分の評価の操作を 1 つのカテゴリで繰り返し行っていると目立つおそれがある。そこで、自分の評価の操作を複数のカテゴリで行い、各カテゴリではあまり目立たないようにするのはではないかと予想した。(予想 2) が正しければ、1 人のユーザが回答者役のユーザ(ユーザ j とする)の評価を高めようとする、ユーザ j は複数のカテゴリで(予想 1)の異常な投稿に回答者として関わることが考えられる。

(予想 3) 1 人のユーザが複数のアカウントを利用して投稿した回答の文体は類似する

(予想 3) が正しければ、1 人のユーザが 2 つのアカウント、ユーザ i とユーザ j を用いて回答を投稿すると、ユーザ i とユーザ j の回答の文体は類似することが考えられる。

そこで、

- (予想 1) と (予想 2) にもとづく異常な投稿行動の検出
- (予想 3) にもとづく文体の類似判定

によって複数のアカウントを利用して質問とその回答を繰り返し投稿し、回答者としての自分の評価を操作しようとした疑いのあるユーザを検出する方法について検討する。

3. 複数のアカウントを用いて自分の評価を操作しようとするユーザの検出

不適切な内容のメッセージが投稿されるのを抑制するため、利用者にユーザ登録することを求める Q&A サイトが多い。しかし、ユーザ登録の情報は、不適切な内容のメッセージを投稿するユーザを検出するには有効であるが、コミュニケーションを巧妙に操作しようとするユーザを検出する手がかりとしては不十分である。そこで、投稿履歴と文体の特徴を手がかりに、

- 異常な投稿を行っている質問者と回答者のユーザペアの検出
- 検出されたユーザペア間での文体の類似判定

を行い、複数のアカウントを用いて自分の評価を操作しようとするユーザを検出する(図-1)。

3.1 異常な投稿を行っているユーザのペアの検出

ユーザ i とユーザ j がそれぞれ質問者と回答者として異常な投稿、すなわち、

1. ユーザ j はユーザ i の質問に対して異常に多くの回答を投稿する
2. ユーザ i はユーザ j から異常に多くの回答を受け取る
3. ユーザ i はユーザ j の回答に対して異常に多くのベストアンサーを与える

を行っていないかどうかを、以下の仮説 QA1、QA2、QA3 および補助仮説 QA3aux を用いて判定する。

仮説QA1 ユーザ j がユーザ i の質問に対して異常に多くの回答を投稿していないならば、ユーザ j はユーザ i の質問に $N_{QA1}(i, j)$ 回の回答を投稿していると期待できる。

$$N_{QA1}(i, j) = P_{QA1}(i) \times ans(j)$$

ここで $ans(j)$ とはユーザ j によって対象カテゴリに投稿された回答の数で、 $P_{QA1}(i)$ とは対象カテゴリに投稿されたすべての質問の中から無作為に 1 つ選んだ質問がユーザ i によって投稿されたものである確率である。Yahoo!知恵袋のユーザは 1 つの質問に対して 1 回しか回答を投稿できないので、 $P_{QA1}(i)$ は

$$P_{QA1}(i) = \frac{qst(i)}{N_{qst}}$$

となる。ここで $qst(i)$ とはユーザ i によって対象カテゴリに投稿された質問の数で、 N_{qst} とは対象カテゴリに投稿された質問の総数である(表-1)。

もしこの仮説が片側二項検定で棄却されれば、ユーザ j はユーザ i の質問に異常に多くの回答を投稿していると判定する。

仮説QA2 ユーザ i がユーザ j から異常に多くの回答を受け取っていないならば、ユーザ i はユーザ j から $N_{QA2}(i, j)$ 件の回答を受け取っていると期待できる。

$$N_{QA2}(i, j) = P_{QA2}(j) \times qst(i)$$

ここで $qst(i)$ とは対象カテゴリにユーザ i が投稿した質問の数である。 $P_{QA2}(j)$ とは、対象カテゴリに質問を 1 つ投稿した場合に、質問を無作為に選んで回答しているユーザ j から回答を受け取る確率である。Yahoo!知恵袋のユーザは 1 つの質問に対して 1 回しか回答を投稿できないので、 $P_{QA2}(j)$ は

$$P_{QA2}(j) = \frac{ans(j)}{N_{qst}}$$

となる。ここで $ans(j)$ とはユーザ j によって対象カテゴリに投稿された回答の数で、 N_{qst} とは対象カテゴリに投稿された質問の総数である(表-1)。

もしこの仮説が片側二項検定で棄却されれば、ユーザ i はユーザ j から異常に多くの回答を受け取っていると判定する。

仮説QA3 ユーザ i がユーザ j の回答に異常に多くのベストアンサーを与えていないならば、ユーザ i はユーザ j の回答に $N_{QA3}(i, j)$ 回ベストアンサーを与えていると期待できる。

$$N_{QA3}(i, j) = P_{QA3}(j) \times f_{QA}(i, j)$$

ここで $f_{QA}(i, j)$ とはユーザ j がユーザ i の質問に投稿した回答の件数で、 $P_{QA3}(j)$ とはユーザ j のベストアンサー率である。

$$P_{QA3}(j) = \frac{bestans(j)}{ans(j)}$$

ここで $ans(j)$ とはユーザ j によって対象カテゴリに投稿された回答の数で、 $bestans(j)$ とは対象カテゴリでユーザ j に与えられたベストアンサーの数である。ただし、

- ユーザ j の回答すべてにベストアンサーが与えられている場合。すなわち、 $ans(j) = bestans(j)$ である場合。
- 補助仮説QA3auxが棄却された場合。すなわち、ユーザ j の回答に異常に多くのベストアンサーが与えられている場合。

以上のいずれかの場合、 $P_{QA3}(j)$ は

$$P_{QA3}(j) = \frac{N_{bestans}}{N_{ans}} = \frac{N_{qst}}{N_{ans}}$$

表-2 入力として与えられた回答の文体が特定のユーザの文体に類似しているか判定するために用いる素性

素性	説明
S1	すべての文に含まれる形態素
S2	すべての文から取り出した文字3-gram
S3	それぞれの文の文頭の1~10文字
S4	それぞれの文の文末の1~10文字

とする。ここで $N_{bestans}$ とは対象カテゴリにおけるベストアンサーの数である。なお、Yahoo!知恵袋のデータでは、それぞれの質問についてベストアンサーが必ず1つあるので、 $N_{bestans}$ は N_{qst} と等しい(表-1)。もしこの仮説が片側二項検定で棄却されれば、ユーザ i はユーザ j の回答に異常に多くのベストアンサーを与えていると判定する。

補助仮説QA3aux ユーザ j の回答に異常に多くのベストアンサーが与えられていないならば、ユーザ j の回答には $N_{QA3aux}(j)$ 件のベストアンサーが与えられていると期待できる。

$$N_{QA3aux}(j) = P_{QA3aux} \times ans(j)$$

P_{QA3aux} とは対象カテゴリにおける平均ベストアンサー率で

$$P_{QA3aux} = \frac{N_{bestans}}{N_{ans}} = \frac{N_{qst}}{N_{ans}}$$

もしこの仮説が片側二項検定で棄却されれば、ユーザ j の回答には異常に多くのベストアンサーが与えられていると判定する。

仮説QA1、QA2、QA3を、質問と回答をやりとりしたすべてのユーザペアに適用して、3つの仮説すべてが棄却されたユーザペアを異常な投稿を行ったユーザペアとして取り出す。

3.2 文体の類似判定

複数のカテゴリで異常な投稿を行っているとして検出されたユーザペアに対して文体の類似判定を行う。文体の類似判定には仮説WSを用いる。仮説WSが棄却されたユーザペアは、

- 複数のカテゴリで異常な投稿を繰り返している
- 文体が類似している

ことから、複数のアカウントを用いて自分の評価を操作している疑いがあると判定する。

仮説WS ユーザ i はユーザ j よりも多くの回答を対象カテゴリに投稿しているとする。また、分類器 $classifier(i)$ は入力された回答の文体がユーザ i の文体に類似しているかどうか判定する分類器とする。このときもし、ユーザ i とユーザ j の文体が類似していないならば、ユーザ j の回答の文体がユーザ i の回答の文体と類似していると分類器 $classifier(i)$ が判定するのは N_{WS} 回であると期待できる。

$$N_{WS} = (1 - P_{WS}) \times m$$

ここで P_{WS} は分類器 $classifier(i)$ の分類精度で、 m は文体の類似判定のために $classifier(i)$ に与えられたユーザ j の回答の件数である。

もしこの仮説が片側二項検定で棄却されれば、ユーザ i とユーザ j の文体は類似していると判定する。

入力された回答の文体が対象ユーザであるユーザ i の文体と類似しているかどうかを判定する分類器 $classifier(i)$ は、ユーザ i の文体に対する最大エントロピー法による学習にもとづいて作成する。

学習に用いるデータは、

- 対象ユーザが対象カテゴリに投稿した m 件の回答から無作為に取り出した n 件の回答
- 対象ユーザ以外のユーザが対象カテゴリに投稿した回答から無作為に取り出した n 件の回答

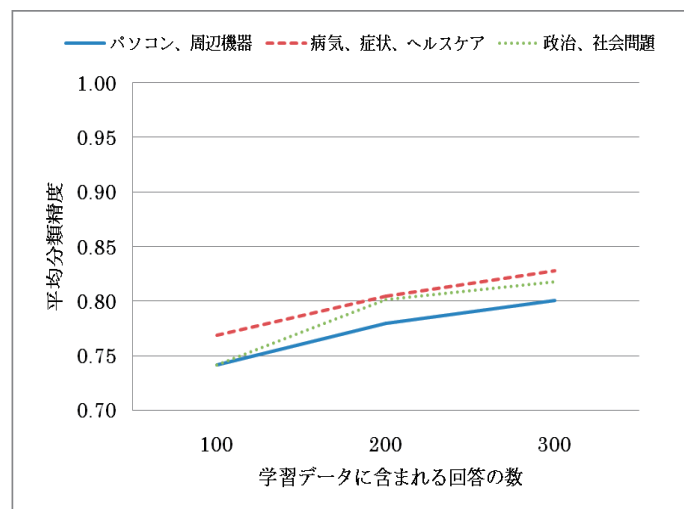


図-2 対象カテゴリで質問を10件以上、回答を160件以上投稿したユーザ30人の回答を対象に作成したユーザ分類器の平均分類精度。学習には、100件(対象ユーザ50件とその他50件)、200件(対象ユーザ100件とその他100件)、300件(対象ユーザ150件とその他150件)の回答からなる学習データを用いた。

の合計 $2n$ 件の回答から構成されている。

学習に用いる素性を表-2に示す。 $S1$ は回答を形態素解析した結果得られた形態素を素性に利用している。 $S2$ で文字 n -gram として 3-gram を使用しているのは、文字 n -gram のうち日本語には 3-gram が効果的であると報告されているからである[9]。 $S3$ と $S4$ でそれぞれの文の文頭および文末の1~10文字を素性としているのは、回答者それぞれの特徴的な表現が文頭および文末に現れることが多いと考えたからである。

分類器の分類精度 P_{WS} を測定するため、Yahoo!知恵袋の「パソコン、周辺機器」「病気、症状、ヘルスケア」「政治、社会問題」のカテゴリを対象に実験を行った。各カテゴリにおける対象ユーザは、質問投稿数が10件以上、回答投稿数が160件以上のユーザ30人とした。実験では最初に、各ユーザの回答から10件を無作為に取り出し、評価用データとした。次に、

- 対象ユーザが対象カテゴリに投稿した回答(評価用データとして取り出した回答10件は除く)から無作為に取り出した n 件の回答
- 対象ユーザ以外のユーザが対象カテゴリに投稿した回答から無作為に取り出した n 件の回答

の合計 $2n$ 件の回答 ($n = 50, 100, 150$) から構成される3種類の学習用データを作成した。この学習データに対して最大エントロピー法による学習を maxent^9 を用いて行った。また、形態素解析には Mecab^{10} を用いた。作成した分類器の分類精度を図-2に示す。図-2に示すように、300件の回答(対象ユーザの回答150件を含む)から構成される学習データを用いれば、どのカテゴリでも80%以上の精度で対象ユーザの回答であるかどうかを判定できる。表-2の素性の中では、それぞれの文の文頭および文末の1~10文字($S3$ と $S4$)が有効であった。

4. 実験と検討

提案手法の有効性を評価するために、Yahoo!知恵袋(286カテゴリ)に回答を投稿したすべてのユーザ(表-1)を対象に、複数のアカウントを利用して1人で質問とその回答を繰り返し投稿し、回答者としての自分の評価を操作した疑いがあるユーザを検出する。

表-3に実験で用いた有意水準を示す。表-3に示すように、仮説 $QA1$ 、 $QA2$ 、 $QA3$ 、 $QA3aux$ の検定に用いた有意水準 0.000005 は一般的に用いられる有意水準に比べて極めて小さい。これは、極めて異常な投稿を検出するためである。表-4に仮説 WS の検定で用いるユーザ分類器の分類精度 P_{WS} を示す。分類精度 P_{WS} には、3.2節で測定した「パソコン、周辺機器」「病気、症状、ヘルスケア」「政治、社会問題」カテゴリにおける分類精度(図-2)の平均値を用いる。 P_{WS} の値は、表-4に示すように、学習に用いる回答の数によって変化する。

⁹ <http://mastarpj.nict.go.jp/mutiyama/software/maxent/>

¹⁰ <http://mecab.sourceforge.net/>

表-3 実験で用いた有意水準

仮説	有意水準
QA1,QA2,QA3,QA3aux	0.000005
WS	0.005

表-4 ユーザ分類器の分類精度 P_{WS}

学習に用いた 回答の件数	分類精度 P_{WS}
～ 99	－
100 ～ 199	0.753
200 ～ 299	0.796
300 ～	0.816

表-5 異常な投稿を行っているユーザペアの文体の類似判定の結果

(a) すべてのカテゴリ (286 カテゴリ)					
	type 1	type 2	type 3	type 4	合計
type a	50	23	5	57	135
type b	32	41	0	34	107

(b) 「パソコン、周辺機器」カテゴリ					
	type 1	type 2	type 3	type 4	合計
type a	2	2	0	3	7
type b	0	0	0	0	0

(c) 「政治、社会問題」カテゴリ					
	type 1	type 2	type 3	type 4	合計
type a	3	3	0	1	7
type b	1	3	0	0	4

実験ではまず、異常な投稿を行ったユーザペアを(予想 1)にもとづいて 387 組検出した。この 387 組のユーザペアから、その回答者が複数のカテゴリで異常な投稿に回答者として関わっているユーザペアを(予想 2)にもとづいて 242 組取り出した。この 242 組のユーザペアは以下の 2 種類に分類できる。

type a 複数のカテゴリで異常な投稿を行っているユーザペア。

type b 1 つのカテゴリでだけ異常な投稿を行っているユーザペア。ただし、その回答者は複数のカテゴリで異常な投稿に回答者として関わっている。

また、文体の類似判定結果によって以下の 4 種類にも分類できる。その結果を表-5(a)に示す。

type 1 仮説 WS が棄却され、類似する文体をもつと判定されたユーザペア

type 2 仮説 WS が棄却されず、類似する文体をもつとは判定されなかったユーザペア

type 3 どのユーザも回答投稿数が 50 件未満であるため、学習データが小さすぎるとして分類器を作成せず、文体の類似判定を行わなかったユーザペア

type 4 一方のユーザの回答投稿数が 4 件未満であるため、仮説 WS が棄却される可能性がないので、文体の類似判定を行わなかったユーザペア

この 242 組のユーザペアすべてについて質問者と回答者が同一人物であるか検討するのはむずかしい。そこで、「パソコン、周辺機器」「政治、社会問題」カテゴリで異常な投稿を行った 18 組のユーザペアについて質問者と回答者が同一人物であるか詳しく検討する。

「パソコン、周辺機器」カテゴリと「政治、社会問題」カテゴリで異常な投稿を行ったユーザペアの文体の類似判定の結果を表-5(b)と表-5(c)に示す。さらに、図-3 に「パソコン、周辺機器」「政治、社会問題」カテゴリで異常な投稿を行ったユーザペアの

- 文体の類似判定の結果
- 同一人物による投稿であるかどうかの検討結果

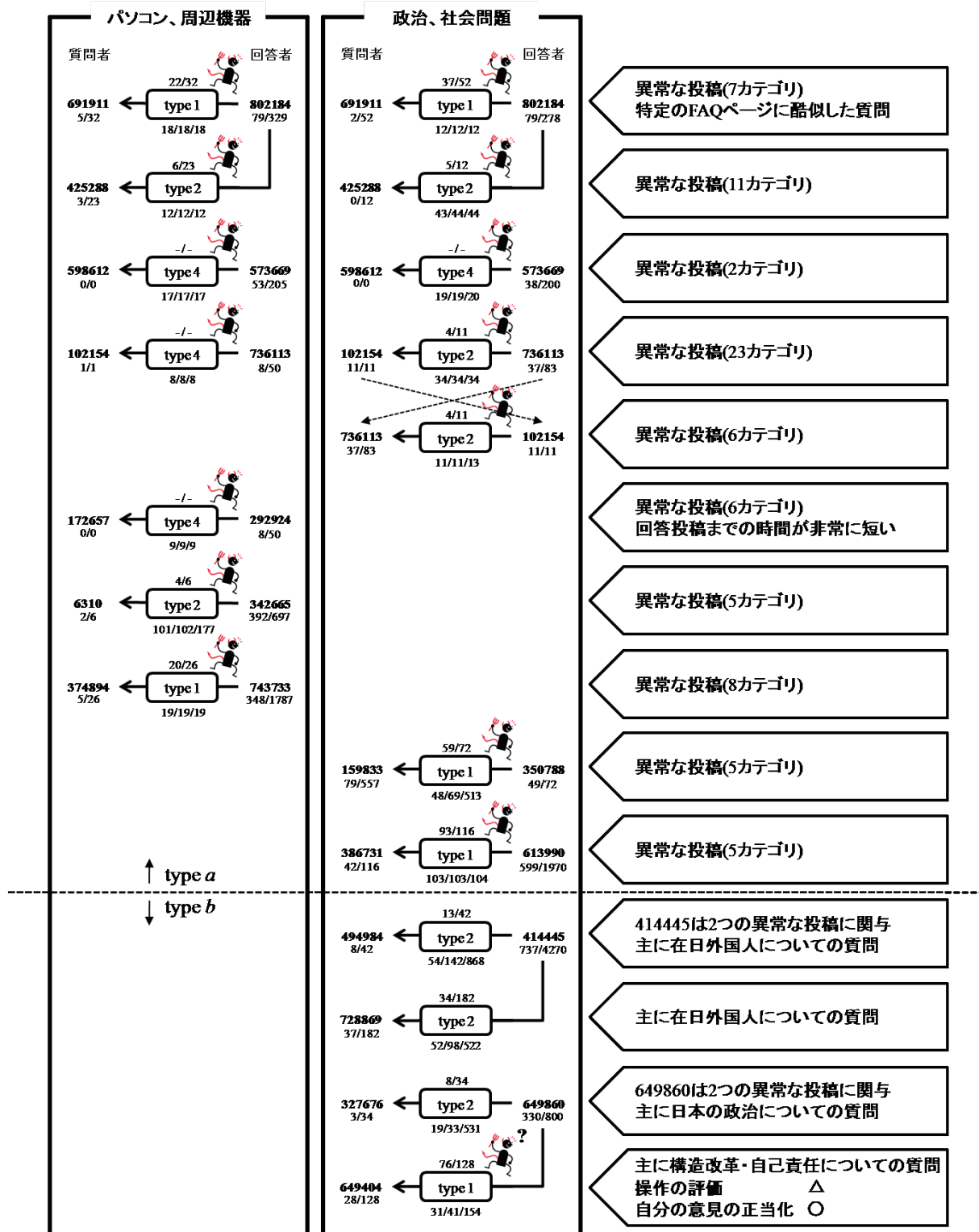


図-3 「パソコン、周辺機器」カテゴリ(7組)と「政治、社会問題」カテゴリ(11組)で異常投稿を行った18組のユーザペアに対する文体の類似判定の結果と質問者と回答者が同一人物であるかの判定結果

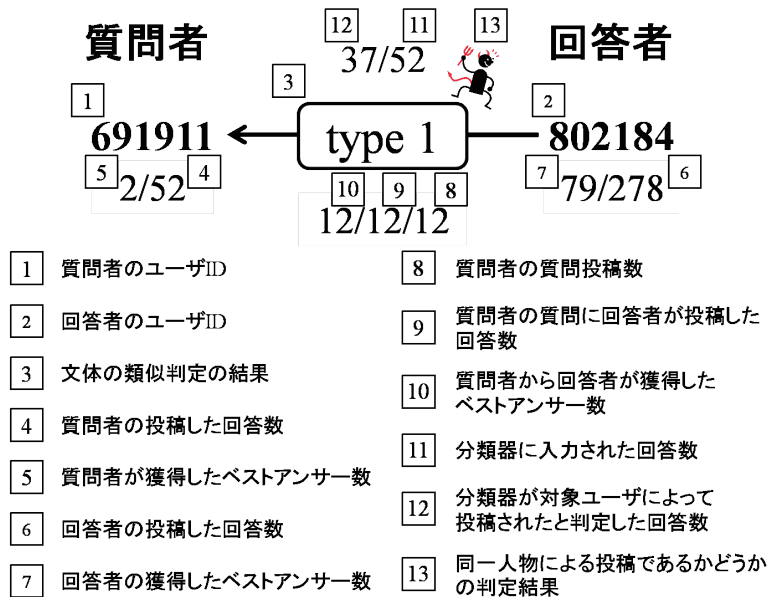


図-4 図-3における数字・記号などの意味

を示す。図-3における数字・記号などの意味を図-4に示す。例えば、図-3の691911と802184はそれぞれ異常な投稿を行った質問者と回答者のユーザIDで、ユーザ691911が「政治、社会問題」カテゴリに投稿した12件の質問すべてにユーザ802184は回答を投稿し、その12件の回答すべてにユーザ691911はベストアンサーを与えている。また、ユーザ691911の投稿した52件の回答のうち37件がユーザ802184のユーザ分類器によってユーザ802184の文体と類似していると判定され、仮説WSによる類似判定の結果、このユーザペア(691911 ← 802184)¹¹はtype 1(文体が類似している)と判定された。そして、ユーザ691911とユーザ802184の投稿を調べ、このユーザペア(691911 ← 802184)では質問者と回答者が同一人物であると判定した。

最初に、「パソコン、周辺機器」「政治、社会問題」カテゴリで異常な投稿を行ったtype aのユーザペア14組について、質問者と回答者が同一人物である可能性について検討した。その結果、type 1以外のユーザペアも含め、14組すべてについて質問者と回答者が同一人物であることが疑われるケースと判定した。このように判定した理由のうち、注目すべきものを以下に示す。

- あるユーザのさまざまな質問に同一の回答者が複数のカテゴリで頻繁に回答するのは不自然である。
- 質問者が投稿した回答を調べると、回答者の意見とよく一致する。
- ユーザペア(691911 ← 802184)の質問者(ユーザ691911)は、あるFAQページにある質問と酷似した内容の質問を「パソコン、周辺機器」カテゴリに繰り返し投稿している。これらの質問に対して回答者(ユーザ802184)は、そのFAQページのURLを示して参考にするように勧める回答を繰り返し投稿している。これらの投稿は、質問と回答を作成する手間を省くために行っているのではないかと考えられる。
- ユーザペア(102154 ← 736113)とユーザペア(736113 ← 102154)は質問者と回答者が逆であるが、どちらのユーザペアも複数のカテゴリで異常な投稿を行っている。
- ユーザペア(172657 ← 292924)では回答が投稿されるまでの時間が非常に短い。これについては本章の最後で説明する。

なお、質問者と回答者が同一人物であることが疑われるケースと判定したこのtype aのユーザペア14組の回答者が投稿した回答は、「パソコン、周辺機器」カテゴリの全回答の0.68%、「政治、社会問題」カテゴリの全回答の0.65%をそれぞれ占める。

次に、「政治、社会問題」カテゴリで異常な投稿を行ったtype bのユーザペア4組について検討する。これらのユーザペアの質問者と回答者は、特定のトピックに強い関心をもっている。例えば、ユーザペア(494984 ← 414445)とユーザペア(728869 ← 414445)は在日外国人に関する質問と回答を繰り返している。しかし、(494984 ← 414445)、(728869 ← 414445)、(327676 ← 649860)の3つのユーザペアでは、質問者と回答者の意見はしばしば対立している。このため、この3組のユーザペアについては質問者と回答者は同一人物ではないと考えられ

¹¹ 以後、異常な投稿を行っているユーザペアを(質問者のユーザID ← 回答者のユーザID)で表現する。

る。一方、ユーザペア (649404 ← 649860) では質問者と回答者の意見は「低所得は自己責任」で一致していた。しかし、(予想 1)に反して、ユーザ 649404 はユーザ 649860 の回答にベストアンサーを与えないことがある。ただし、その場合、ユーザ 649404 は自分と同じ意見の投稿にベストアンサーを与えていた。したがって、このユーザは、自分の評価を操作することよりも、反対意見の投稿をベストアンサーに選びたくないことが目的と考えられる。

「パソコン、周辺機器」と「政治、社会問題」カテゴリでの調査から、複数カテゴリで異常な投稿を行う type a のユーザペア 135 件の多くが質問者と回答者が同一人物である疑いが強いと考えられる。一方、(予想 3)にもとづく文体の類似判定の有効性は限定的で、図-3 に示すように疑わしいユーザペアでも type 2 に分類されることが多い。現在、(予想 3)について見落としがないか検討しているが、図-3 で質問者の回答投稿数が少なく type 4 に分類された 4 組の疑わしいユーザペアに注目している。(予想 1)が正しいなら、複数のアカウントを利用して自分の評価を操作しようとするユーザはできるだけ少ない投稿回数でその目的を達成しようとする。したがって、質問者役のユーザが投稿する回答は少なくなると考えられる。実際、表-5 に示すように、type a のユーザペア 135 組中、57 組が type 4 に分類されている。このため、文体の類似判定以外で疑わしいユーザペアを検出する方法を検討することが重要である。例えば、回答が投稿されるまでの時間は手がかりとして有望である。例えば、「パソコン、周辺機器」カテゴリで type 4 に分類されているユーザペア (172657 ← 292924) では、ユーザ 172657 の質問にユーザ 292924 が回答を投稿するまでの時間は非常に短く、最長で 46 秒であった。特に、400 回以上キー操作が必要な回答をわずか 28 秒で投稿している例もあった。このユーザは、回答をあらかじめ準備してから質問を投稿していると考えられる。

5. おわりに

本研究では、複数のアカウントを利用して 1 人で Q&A サイトに質問とその回答を繰り返し投稿し、自分の評価を操作しようとするユーザを、異常な投稿の検出と文体の類似判定を用いて検出する方法を提案した。そして、質問者と回答者が同一人物であることが疑われるユーザペアを 135 組検出し、それらの検出には異常な投稿の検出が有効であることを示した。質問者と回答者が同一人物であることが疑われるユーザペアの例は、複数のアカウントを用いてコミュニティサイトでのコミュニケーションを操作しようとするユーザの目的や方法について調査するのに役立つと考えられる。

【謝辞】本研究の一部は、日本学術振興会科学研究費補助金基盤(C)「心豊かなコミュニケーションを促進する質問作成支援システムの作成」(課題番号 20500106)および挑戦的萌芽研究「質問の多様な形態と機能を明らかにし利用する「質問工学」の創成」(課題番号 22650021)の助成を受けて行われたものです。

【参考文献】

- [1] 三浦, 川浦, 地福, 大瀧, 岡本: 知識共有コミュニティを創り出す人たち, 人工知能学会全国大会 (第 20 回), 3D3-1, (2006).
- [2] de Vel, Anderson, Corney, and Mohay: Mining e-mail content for author identification forensics, ACM SIGMOD Record, 30(4), (2001).
- [3] Koppel, Argamon, and Shimoni: Automatically Categorizing Written Text by Gender, Literary Linguistic and Computing, 17(4), (2002).
- [4] Corney, de Vel, Anderson, and Mohay: Gender-Preferential Text Mining of E-mail Discourse, ACSAC 2002, (2002).
- [5] Argamon, Saric, and Stein: Style mining of electronic messages for multiple authorship discrimination: first results, 9th ACM SIGKDD, (2003).
- [6] Zheng, Li, Chen, and Huang: A Framework of Authorship Identification for Online Messages: Writing Style Features and Classification Techniques, Journal of the American Society for Information Science and Technology, 57(3), (2006).
- [7] 坪井, 松本: Authorship Identification for Heterogeneous Documents, 情報処理学会研究報告, 2002-NL-148, (2002).
- [8] 折田, 江木: リンク不能性および一覧性の観点による匿名性の分類, 情報処理学会研究報告, 2007-EIP-37, (2007).
- [9] 小高, 村田, 高, 諏訪, 白井, 高橋, 黒岩, 小倉: n-gram を用いた学生レポート評価手法の提案, 電子情報通信学会論文誌, Vol. J86-D-I No. 9, (2003).