

Title	Q&Aサイトにおける社会調査型質問への回答者に対する信頼性判断支援システム
Author(s)	王, 曦虹; 小林, 智也; 小倉, 加奈代; 西本, 一志
Citation	情報処理学会研究報告, 2012-HCI-147(18): 1-8
Issue Date	2012-03-22
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/10640
Rights	<p>社団法人 情報処理学会, 王曦虹, 小林智也, 小倉加奈代, 西本一志, 情報処理学会研究報告, 2012-HCI-147(18), 2012, 1-8. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は(社)情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	

Q&A サイトにおける社会調査型質問への回答者 に対する信頼性判断支援システム

王 曦虹[†] 小林智也[†] 小倉加奈代[†] 西本一志[†]

近年 Q&A サイトの利用者が増加している。Q&A サイトの質問は情報検索型と社会調査型に大別される。このうち社会調査型質問には客観的な正解はなく、特定の個人あるいは集団に対してアンケート調査を行うことで回答を得る。このため、質問者は回答の信頼性を判断しにくいという問題点がある。本稿では、質問者が回答の信頼性を判断するための補助材料として、回答者の質問・回答履歴から、今回回答しようとしている質問と類似した質問とそれへの回答を抽出し、これを質問者に提示する。これによって、質問者が回答者の態度や性格などの特性を推測可能とする手法を提案する。Yahoo! 知恵袋のデータを用いてシステムの評価を行い、一定の有用性を確認した。

An assistant system for judging reliability of responses to social survey questions in a Q&A website

Xihong Wang[†] Tomoya Kobayashi[†]
Kanayo Ogura[†] Kazushi Nishimoto[†]

Recently, Q&A sites have been widely used. Questions of the Q&A sites can be classified into two types: an information-seeking type and a social-survey type. Since there are no objectively correct answers for the social-survey type questions, questionnaires are usually used to obtain answers from responders. However, it is difficult for a questioner to judge credibility of the answers. For supporting the judgment, we extract questions similar to the questioner's question from the responders' histories of questions and answers as well as answers to the extracted questions. By providing them to the questioner, it is expected that the questioner becomes able to infer each responder's attitude and sense of values. We conducted pilot studies using Yahoo! Chiebukuro data and confirmed basic efficiency of the proposed method.

1. はじめに

近年、Q&A サイトと呼ばれる質問回答サイトの利用者が増加している。Q&A サイトというのは会員同士が、お互いの質問に答え、疑問を解決するウェブサイトのことである[1]。日本で代表的な Q&A サイトとしては、登録者数第 1 位の Yahoo! 知恵袋や第 2 位の OKWave などが挙げられる。栗山らは、Q&A サイトにおける質問には情報検索型質問と社会調査型質問の 2 種類があることを指摘している[2]。情報検索型質問とは、サーチエンジンや図書館レファレンスサービスを利用して回答を探すことが可能な内容に関する質問である。社会調査型質問とは、客観的な唯一の正解が存在せず、特定の個人あるいは集団に対してアンケート調査を行うことで各回答者の主観に基づく回答を得るような質問である。

本稿では、社会調査型質問を対象とし、得られた個々の回答の良否を判断する際の助けとなる情報を提供するシステムを提案する。社会調査型質問に対しては、回答者それぞれが大きく異なった内容の回答を寄せる。各回答の正誤を判断するための客観的な基準が存在しないため、質問者は、どの回答を良い回答として受け入れるかを自分の主観に基づき判断せざるを得ない。また、多くの Q&A サイトでは、寄せられた回答を質問者が評価し、ベストアンサーを選んで優れた回答者を報賞することが求められている。社会調査型質問の場合は、ベストアンサーの選定も主観的に行わざるを得ない。

社会調査型質問においては、ある回答の信頼性の判断を行う際に、その回答の内容だけに基いて判断するのが難しいと筆者らは考える。その回答を行った回答者の特性に関する情報も加味する必要がある。たとえば、ある回答者があちこちで内容の一貫性がない回答をしていた場合、その回答者による回答の信頼性は低いのではないかと思われる。あるいは、非常に似た回答が 2 つ寄せられ、いずれかをベストアンサーとして選定する場合、質問者はこの分野に経験が豊富な回答者による回答をベストアンサーとしたいと考えることも想定される。

本稿では、社会調査型質問における各回答の良否判断を支援するために、各回答者の質問履歴および回答履歴を用いて、質問者が各回答者の特性について判断するための参考情報を提供するシステムを提案する。Yahoo! 知恵袋のデータを利用して、提案システムを使用した場合と使用しなかった場合とを比較することにより、提案手法の有効性を評価する。以下、第 2 章では Q&A サイトについての関連研究を述べる。第 3 章では、本研究で提案する手法及び理由について説明する。第 4 章では、提案シ

[†] 北陸先端科学技術大学院大学 知識科学研究科

School of Knowledge Science, Japan Advanced Institute of Science and Technology

[†] 北陸先端科学技術大学院大学 ライフスタイルデザイン研究センター

Research Center for Innovative Lifestyle Design, Japan Advanced Institute of Science and Technology

システムの概要について述べる。第5章では、提案したシステムにおける類似度判断の妥当性を評価する予備実験とその結果について述べる。第6章では、提案手法の有効性の評価実験について述べる。第7章はまとめである。

2. 関連研究

高田ら[3]は、Web情報を用いてコンテンツを補完する手法を提案した。情報検索型質問の場合は、回答の信頼性を判断するための情報をWebから収集し補完する。社会調査型質問の場合は、今の回答において不足している回答情報をほかの類似Q&Aコンテンツから収集し補完する。この手法は、得られた回答に情報を追加して回答の内容を改変することによって回答の質を向上させることを狙っているものであり、元の回答自体の良否判定を支援するものではない。

瀧ら[4]は、Q&Aコミュニティ内で正しい回答を知らない質問者が正しい回答を判断しようとすることを支援する信頼性指標を提案した。質問者が正しい回答かどうかを判断する際に参考にする情報を増やすために、回答内容だけでなく、回答の外部参照、質問のタイミングなどの付随情報をも手がかりとして利用していることを指摘した。この研究では、元の回答自体の良否判定を支援している点で、我々の研究と目的を同じくしている。ただし、そのために利用する情報として、我々の研究では回答者の質問・回答履歴を用いている点が異なっている。

佐々木ら[5]は、回答の根拠が提示されないことが、回答として提示された行動が適切かどうかの判断を困難にしていることを指摘し、原因や理由を回答する why 型 Q&A を応用し、取るべき行動と理由を合わせて提示する Q&A を提案した。これにより、ただ取るべき行動のみを示されるよりは、回答の良否判定のための判断材料が増えることは間違いない。しかしながら、特に社会調査型質問の場合、行動と根拠の両方が常に提示されたとしても、その回答の良否を最終的に判断するためには、その回答者の特性に関する情報が必要であると考えられる。

3. 提案手法

本研究の基本的な発想は、社会調査型質問の場合、質問者は、自分の質問と同じ分野に関する回答者の特性が分かれば、回答の信頼性を判定できる可能性が高いのではないかと考える。ゆえに、各回答者が持つ特性を質問者が把握することを支援できれば、各回答の信頼性判定の一助となることが期待される。ただし、回答者の特性はきわめて個人差も大きく多様であり、しかも一個人がもつ特性に関しても、しばしば非論理的であったり、一貫性がなかったり、曖昧であったりする。このため、特性を機械処理によって抽出することは現実的では無いと思われる。

そこで本研究では、質問者（すなわち提案システムのユーザ）に対して、回答者の

特性そのものを提示するのではなく、特性を推定するための素材となる情報を提示し、推定はユーザにまかせる手段をとる。特性推定のための情報源として、回答者の過去の質問履歴と回答履歴を利用する。ただし、すべての履歴を用いるのではなく、質問で触れたことに関連性がある履歴のみを抽出・提示する手段をとる。例えば、回答者の回答の一貫性という特性について、すべての履歴から判断することは難しいと考える。ある回答者による恋愛問題に対する回答と、消費税問題に対する回答を並べてみても、その回答者が一貫した主張を持っているかどうかを判断することは無理ではないかと考える。しかし、消費税問題に関する回答だけを集めて見てみれば、少なくとも消費税問題に関して一貫性があるかないかの判断は可能だと考える。

参考情報として、具体的には以下に述べる情報を抽出し、これを質問者に提供する。**質問履歴からの参考情報（図1）**

ユーザによる質問と、ある回答者が過去に行ったすべての質問とを比較し、一定以上の類似度（類似度の算出方法は後述）を持つ類似質問を抽出する。ついで、抽出された類似質問に対して寄せられた回答のうち、その回答者がベストアンサーと評価した回答を抽出する。このベストアンサーの選定には、その回答者の考え方が反映されているはずである。そこで、こうして得られた類似質問とベストアンサーの組を、その回答者の特性判断のための素材情報としてユーザに提示する。

回答履歴からの参考情報（図2）

ユーザによる質問と、ある回答者が過去に回答したすべての質問とを比較し、一定以上の類似度（類似度の算出方法は後述）を持つ類似質問を、その回答者による回答

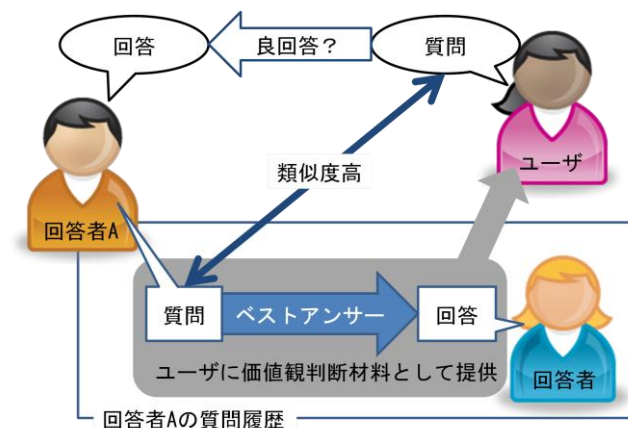


図1 質問履歴からの情報

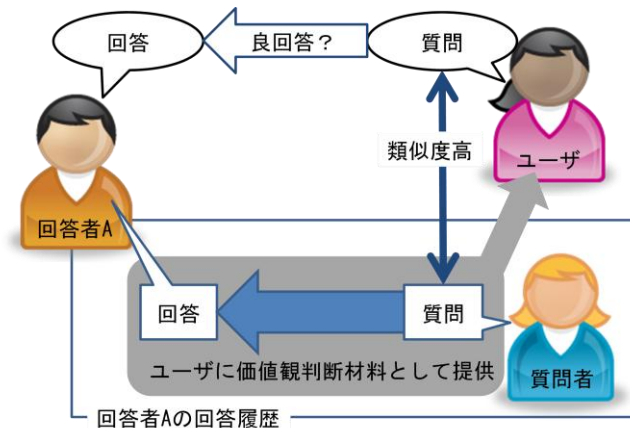


図 3 回答履歴からの情報

と合わせて抽出する。この回答には、その回答者の特性が反映されているはずである。そこで、こうして得られた類似質問と回答の組を、その回答者の特性判断のための素材情報としてユーザに提示する。

以上により、ユーザによる質問に関する分野に対して、各回答者がどのような価値観を持っているかを推定するための有益な情報を提供することができると考えられる。

4. システム概要

本システムは、ウェブブラウザ上で動作するシステムとして構築し、Mozilla Firefox のアドオン機能として動作するように開発した。クライアント側システムは Greasemonkey 上で動作する JavaScript、サーバ側システムは Apache ウェブサーバ上で動作する PHP で実装した。システム構成の概要を図 2 に示す。質問者が Firefox で Q&A コンテンツを閲覧すると、Greasemonkey が各回答の回答者のすぐ下にシステムを呼び出すためのボタンを追加する。追加されたボタンを押すと質問文と回答者リストを取得して、サーバ側に送信する。サーバ側では Yahoo! テキスト解析 API を用い、受信した質問文のキーワードを抽出する。また Yahoo! 知恵袋の情報をダウンロードし、回答者の回答履歴および質問履歴中の質問文をダウンロードする。ダウンロードした質問文も同様にキーワードを抽出する。質問文のキーワードと履歴中質問文のキーワードの類似度を計算し、類似質問を抽出して HTML ページを生成し

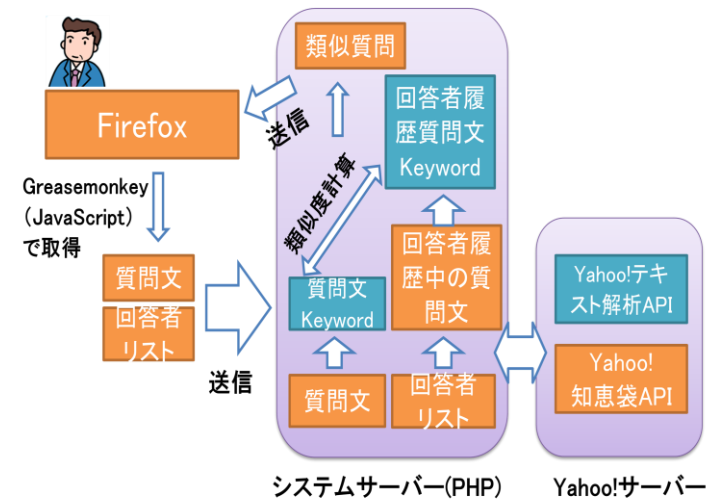


図 2 システム概要

Firefox に参考情報として表示する。

キーワードの抽出方法は、質問文を形態素解析し、名詞と動詞をキーワードとして抽出することに加え、カテゴリ情報も抽出する。類似度の計算方法にはベクトル空間法を用いた。キーワード毎に 1 次元とし、キーワードの出現回数をベクトルの重みとする質問文キーワードベクトルを作り、以下の式によって 2 つの質問の質問文キーワードベクトル \vec{Q}_1, \vec{Q}_2 の成す角の余弦値 $\cos\theta$ を計算し類似度 $\text{sim}(\vec{Q}_1, \vec{Q}_2)$ とする [6].

$$\text{sim}(\vec{Q}_1, \vec{Q}_2) = \cos\theta = \frac{\vec{Q}_1 \cdot \vec{Q}_2}{|\vec{Q}_1| |\vec{Q}_2|}$$

システムのユーザ・インターフェースを図 4 に示す。回答者アイコンの下には「参考情報」というボタンが表示される。質問者がボタンを押すと、別のページで類似質問の参考情報が提示される。質問履歴を用いた情報提供の場合は、この回答者が以前質問した履歴の中で、今回回答しようとしている質問との類似度がゼロでない質問に対して、この回答者が選んだベストアンサーを、質問の類似度の高い順に提示する。回答履歴を用いた情報提供の場合は、この回答者が以前回答した履歴の中で、今回回答しようとしている質問との類似度がゼロでない質問に対するこの回答者の回答を、類似度の高い順に提示する。

5. 類似度評価に関する予備実験

5.1 手順

4章で説明した手法によって求めた質問の類似度が、実際の人間が評価する類似度と一致しているかどうかを検証するため、アンケートを実施した。被験者は、著者が在籍する大学院の学生5名（うち男性4人女性1人）である。

実験に使用したQ&Aコンテンツには、Yahoo!知恵袋に投稿された表1の5つのカテゴリから15個のQ&Aコンテンツを選んだ。これら15個の各質問（以下、「元質問」とする）に対して回答した回答者が、過去に回答した別の質問および自らが質問者となって行った質問（以下、これらをまとめて「過去質問」とする）を集める。こうして得られた、元質問と過去質問の組、合計15組を被験者に提示し、各組に関して、その組に含まれる元質問と各過去質問とを比較し、1:非常に似ていない 2:似ていない 3:どちらかというのと似ていない 4:どちらかというのと似ている 5:似ている 6:非常に似ている 0:分からない、の7段階評価で評価するアンケートに回答してもらった。一方、4章で述べた方法で、各組に関して、その組に含まれる元質問と各過去質問との類似度を計算する。これを、上述の被験者による評価結果と比較し、両者の相関関係を分析する。なお、15個のQ&Aコンテンツに対する回答者の総数は27人、過去質問の総数は616個であった。

5.2 結果

実験期間にも質問履歴から質問を削除した回答者が存在したため、被験者間で読むことのできる質問に差違があり、最終的に100ぐらいの質問が除外された。残った516個の質問について、システムが算出した類似度(0~1)と被験者の評定値(1~6)とに関する相関係数を求めた。相関係数は0.328で、弱い正の相関があり、検定の結果5%水準で有意であった。この結果から、4章で採用した質問の類似度計算手法により、人間による評価と相関のある類似度が得られることがわかった。

6. 評価実験

6.1 被験者が質問者である場合

提案手法の有効性に関する評価のために、4章で述べたシステムを使用した場合としない場合の比較実験を実施した。被験者は、著者の在籍する大学院の学生10名（うち男性8人、女性2人）である。

まず、実験の準備として、被験者にYahoo!知恵袋で質問を投稿してもらった。投稿質問の種類は、どうすればいいかわからない悩みごとに限られた。10人の被験者全員に、1人につき質問3問を投稿してもらい、全部30件の質問を得た。回答の収集期間は投稿してから1週間である（Yahoo!知恵袋における回答受付の期間は1週間、

表1 選んだQAコンテンツ

1	エンターテインメントと趣味	①話題の人物, ②アニメ, ③携帯型ゲーム全般
2	暮らしと生活ガイド	④レシピ, ⑤洗濯, クリーニング, ⑥引っ越し
3	健康美容とファッション	⑦花粉症, ⑧ダイエット, ⑨ファッション
4	生き方と恋愛, 人間関係の悩み	⑩恋愛相談, ⑪友人, ⑫一人暮らし
5	子育てと学校	⑬子育ての悩み, ⑭大学院, ⑮大学の受験

回答受付締め切りになってもベストアンサーを選ばなかったら、投票に移る)。できる限り多くの回答を集めるため、回答受付締め切りまで回答を収集した。

6.1.1 実験手順

本実験はシステム利用前とシステム利用後の2部分に分けて実施された。まず被験者は、システムを使用せずに回答を閲覧し、ベストアンサーを選択する。選択が終わったら、システム利用前アンケートに回答する。アンケートでは、ベストアンサーの選択理由とその確信度(5段階評価)、ベストアンサー候補の数、および候補には挙げたがベストアンサーとしなかった回答に関して、それらをベストアンサーとしなかった理由を尋ねた。次いで、提案システムを利用して参考情報を提示する状態で、被験者はもう一度ベストアンサーを選択する。その後、システム利用後のアンケートに回答する。アンケートには、システム利用前と同じ質問に加えて、提示された参考情報からどの程度回答者の考え方を理解できたか、回答者の回答履歴にどの程度一貫性があると思ったか、ベストアンサーを決める際にどの程度参考情報を参考にしたかについて、それぞれ5段階(いずれについても、1が最も低い評価値であり、5が最も高い評価値とした)で評価してもらった。

提案システムの利用後、質問者が利用前に選んでいたベストアンサーを変更したり、ベストアンサーに対する確信度が高くなったりすれば、それは提案システムが提供した参考情報が有効に影響した結果であると見なすことができよう。

6.1.2 結果

30問の質問に対して、平均の回答数は2.5件であった。この内、回答数が0件のものが2問あり、ベストアンサー候補数が0件のもの(質問の言葉が足りなかったので、得られた回答は質問者の本意と違ったもの)も1問あった。よって、以下の分析に使える質問は27問であった。

表 2 システム利用前と比べたシステム利用後の結果

	変わった	変わらなかった
ベストアンサー候補者数	3 件 (11%)	24 件 (89%)
ベストアンサー回答者	6 件 (22%)	21 件 (78%)
ベストアンサー候補点数	39 名 (76%)	12 名 (24%)

表 3 参考情報について評価の平均値

被験者の評価	平均
回答者の考え方に対する理解度	3.82
回答者履歴の一貫性	4.02
ベストアンサーの回答者	4.07
ベストアンサーではない回答者	4.00
質問履歴の参考度	2.82
回答履歴の参考度	3.98

表 2 は、システム利用前と比べて、システム利用後に変わったことと変わらなかったことのまとめである。システム利用前と利用後の変化から見れば、システム利用後にベストアンサー候補者の人数がシステム利用前より減ったのは 3 件、変わらなかったのは 24 件であった。ただし、変わらなかった 24 件のうち、回答が 1 つしかなかったものが 8 件あった。これらについては、システム利用前も利用後もベストアンサー候補が 1 人しかいなかったため、変わらなかったのではなく、変えようが無かったものである。ゆえに、現実的に変わらなかった件数は 16 件であった。また、候補数が増加した事例はなかった。このように、ベストアンサー候補数は、システムの利用によって減少する傾向があることがわかった。この結果は、提案システムが提供する参考情報によって、ベストアンサー選択時の迷いがある程度解消され、結果としてベストアンサー候補数がシステム利用前より減少した可能性を示唆している。

システム利用後にベストアンサーが変わったのは 6 件、変わらなかったのは 21 件であった。変わらなかった 21 件のベストアンサーについて、システム利用後に行った被験者アンケートから得られた選定理由は、以下の通りである：

- 【システム利用前と同じ理由】 4 件
- 【類似質問の履歴が多くて、経験豊富で信頼できる】 8 件
- 【履歴を読んだら、誰に対しても丁寧に真面目に答えている】 7 件
- 【解決案を述べる回答履歴が多い】 3 件
- 【オリジナル回答をしていた】 1 件
- 【考え方が一緒】 1 件

● 【前向きな回答を多くしていた】 1 件

システム利用前はベストアンサーとして選定されていたにもかかわらず、システム利用後にベストアンサーとして選定することをやめた回答が 2 件あった。その理由は、「似たような質問に対して全部同じ答えを書いていたから」と「ネットで相手を悪く言う書き込み（質問が多少あり）感情をすぐ言葉に出す所があって、怖いのです」というものであった。

ベストアンサー候補に挙げられたものの、最終的にはベストアンサーとしては選ばれなかった候補者について、やはりシステム利用後にいった被験者アンケートから得られた非選定理由は、以下の通りである：

- 【システム利用前と同じ理由】 4 件
- 【参考情報が少ない】 3 件
- 【一貫性がない】 2 件
- 【参考情報を読んだら、回答者は視野が狭いと感じる】 2 件
- 【履歴で回答者は自分の体験からの回答が少ない】 2 件
- 【参考情報を読んだら、回答者に対する印象が悪い】 2 件
- 【求めている解決方法に関する知識が少ない】 2 件
- 【履歴を読んだら、意見が合わない】 1 件
- 【参考情報には乱暴な言葉を書いた】 1 件
- 【真面目に回答しなかった履歴が多い】 1 件

システム利用後に点数が変わったベストアンサー候補は 39 名で、変わらなかったのは 12 名であった。

このように、システムを利用してもベストアンサーを変更することはあまり多くなかったが、点数と選定・非選定理由はほとんど変わった。システム利用後、ベストアンサーとして選んだ理由・選ばなかった理由を総合的に分析すると、主にこの分野に関する経験が豊富であるか、一貫性があるかなどの、回答者に関する特性にかかわるものが多く、これらの情報によって信頼性を判断でき、ベストアンサーとして選択した、あるいは選択しなかったことへの確信度が高くなっていた。以上の結果から、システムが提示する情報がベストアンサー選択に役に立ったことが分かった。

表 3 はシステム利用後の被験者による参考情報に関する評価の平均値である。表 3 に示す通り、参考情報から被験者が回答者の考え方を理解できた度合いの平均値は 3.82 であった。さらに、被験者が回答者の履歴を読んで、一貫性があると思った評価の平均値は 4.02 であった。その内、ベストアンサーとした回答者に対する一貫性評価の平均値は 4.07 であり、ベストアンサーとされなかった回答者の考え方はより一貫性あると感じられた回答者である傾向が見られた。また、質問履歴からの参考情報が被験者がベストアンサーを決める際に参考になったと思う評価の平均値は 2.82 であったが、回答履歴からの参考情報の場合は、被験者が参考になったと思った評価の平均

値は 3.98 で、質問履歴の場合より評価が高かった。ただしこの結果は、質問履歴の平均数が回答履歴より少なく、0 の場合もいくつかあるなど、履歴の数が少なかったことも原因の 1 つだと思われる。

6.2 被験者が閲覧者である場合

提案システムのユーザは、本来 6.1 の実験で行ったように「質問者」である。しかし、6.1 の実験では十分な量の回答を得られず、平均回答数はわずか 2.5 件であった。このことが、システム利用後にもベストアンサー候補数があまり変化しなかったことの原因だと思われる。

そこで、さらにシステムの有効性を評価するため、この実験では Yahoo!知恵袋にすでに既存している QA コンテンツのうちから、平均回答数 10 件の QA コンテンツを取り出し、これを用いて 6.1 と同様の評価を行った。すなわち、本実験では、被験者は質問者ではなく、閲覧者の立場での実験であると見なせる。

6.2.1 手順

筆者らは「Yahoo! 知恵袋」から 3 つの Q&A コンテンツを選んだ。もっとも社会調査型質問に支配されると考えられる「恋愛相談」カテゴリの「投票受付中の質問」から選んだ QA コンテンツである。平均回答数は 10 件であり、回答内容がそれぞれである Q&A コンテンツを用いた。

被験者は、著者が在籍する大学院の学生 15 名（うち男性 9 人、女性 6 人）である。実験の手順は 6.1 に述べた手順と同じく、システム利用前とシステム利用後の 2 部分にうち分けて実施された。

6.2.2 結果

15 人の被験者全員に、1 人で QA コンテンツ 3 件を評価してもらい、全部 45 件のデータを得た。表 4 に示したシステム利用前と利用後の変化から見れば、システム利用後にベストアンサーとして選びたい回答者の候補数がシステム利用前より減ったのは 19 件 (42%) となり、質問者立場での実験より大きく増加した。このように、ベストアンサー候補数は、システムの利用によって減少することが明らかになった。この結果は、提案システムが提供する参考情報によって、候補選択時の迷いが解消され、

表 4 閲覧者の立場でのシステム利用前と比べたシステム利用後の結果

	変わった	変わらなかった
ベストアンサー候補数	19 件 (42%)	26 件 (58%)
ベストアンサーユーザ名	7 件 (16%)	38 件 (84%)
ベストアンサー候補点数	34 名 (68%)	20 名 (32%)

結果としてベストアンサー候補数がシステム利用前より減少した可能性を示している。

また、システム利用前とシステム利用後のベストアンサーの選択理由を分析した。その結果、システム利用後にベストアンサーとして選んだ理由は主に以下のである。

- 【参考情報が多い】 9 件
- 【考え方が一緒】 8 件
- 【類似質問の履歴が多くて、経験豊富で信頼できる】 5 件
- 【ほかの質問にも真面目に回答した】 4 件
- 【回答者の性格が分かった】 4 件
- 【回答者の持っている知識が分かった】 3 件
- 【回答者の年齢・学歴が分かった】 3 件
- 【回答者の考え方が分かった】 2 件
- 【システム利用前と同じ理由】 3 件
- 【なんとなく】 2 件
- 【一番気に入った】 2 件

システム利用後、ベストアンサー候補に含まれたもののうち、最終的にベストアンサーとして選ばなかった理由は主に以下のである。

- 【回答者の年齢・性格と違うので、信頼しにくい】 7 件
- 【回答が詳しくなくて、説得力がない】 6 件
- 【提供した参考情報が少ない】 5 件
- 【相対的に気に入らなかった】 5 件
- 【参考になった履歴が少ない】 4 件
- 【この分野の経験が無さそうに見える】 4 件
- 【システム利用前と同じ理由】 4 件
- 【考え方が違う】 3 件
- 【回答が普通過ぎ】 3 件

このように、システム利用後、ベストアンサーとして選んだ理由・選ばなかった理由を総合的に分析すると、参考情報を読んだら、主に回答者経験、性格とかの特性が分かって、信頼性を判断でき、ベストアンサー選択に確信が高くなったことが分かった。

7. 考察

本研究では、社会調査型質問への回答を評価するためには、回答だけではなく、回答者の特性に関する情報が必要であると考えた。6 章で述べた実験の結果、類似質問に対する回答者の質問・回答履歴から、被験者は各回答者の特性を推定し、それをも

とに回答の評価を判定することがわかった。特に、ベストアンサー候補者数が減少したという結果から、このような情報がベストアンサーの絞り込みに有効であることが示された。以上から、本研究における「社会調査型質問の場合、質問者は、自分の質問と同じ分野に関する回答者の特性が分かれば、回答の信頼性を判定できる可能性が高いのではないか」という発想は、基本的に正しいということが支持された。

ただし、現状のシステムはまだ十分に質問者に提供する情報を絞り込んでいるとは言いがたい。実験に対する感想の自由記述を見ると、「参考情報が多すぎるので、類似質問に関する情報だけを提供したほうがいい」というコメントが幾つか見られた。現状の類似度計算方法は非常に単純なものであり、人手によって選定された類似質問との相関は十分に高いものとはいえない。この結果、必ずしも類似質問とは言いがたい履歴も抽出された者と思われる。

8. おわりに

本稿では、Q&A サイトにおける社会調査型質問に対する回答を評価するために、回答者の質問履歴と回答履歴を用いて、各回答者の一貫性や経験などの特性を判断するための素材情報を提供するシステムを提案した。回答者の履歴から今回回答しようとしている質問と類似した質問とそれへの回答を抽出し、これを質問者に提示する。提案システムを用いることにより、システムを用いない場合に選択したベストアンサーが覆されることはあまり多く生じなかったが、選択したベストアンサーに対する確信

度が向上する結果が得られた。この結果から、提案手法には一定の有効性があることが明らかとなった。

今後、より高い相関を得られる類似度計算方法を検討したい、また、より広い範囲で提案システムを応用したいと考える。

参考文献

- 1) <http://ja.wikipedia.org/wiki/ナレッジコミュニティ>
- 2) 栗山和子, 神門典子: Q&A サイトにおける質問と回答の分析, 情報処理学会研究報告, (2009).
- 3) 高田夏希, 山本裕補, 小山聡, 田中克己: 質問応答コンテンツに対する Web による回答補完, DEIM Forum 2009 C4-6 (2009).
- 4) 瀧寛文, 森崎修司, 大平雅雄, 松本健一: Q&A コミュニティを対象とした回答の信頼性指標構築に向けた分析, 情報社会学会誌, Vol.4, No.1, pp.49-58 (2009).
- 5) 佐々木智, 藤井敦: 取るべき行動と理由を提示するヘルプデスク指向の質問応答システム, (2010).
- 6) <http://www.gifu-nct.ac.jp/elec/deguchi/sotsuron/havashi/node20.html>