

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

An Analysis toward Constructing Answer Trust Metrics for Q&A Community

瀧 寛文(たき ひろふみ・Hirofumi Taki)¹、森崎 修司(もりさき しゅうじ・Shuji Morisaki)²、
大平 雅雄 (おおひら まさお・Masao Ohira)³、松本 健一(まつもと けんいち・Kenichi Matsumoto)⁴

¹奈良先端科学技術大学院大学情報科学研究科 博士前期課程

²奈良先端科学技術大学院大学情報科学研究科 助教

³奈良先端科学技術大学院大学情報科学研究科 助教

⁴奈良先端科学技術大学院大学情報科学研究科 教授

[Abstract]

In a Q&A community, where anyone can post questions and answers via the Web, there are few clues for a questioner who lacks appropriate knowledge to select a right answer. This study reports an analysis conducted to construct the trust indexes for a Q&A community. From the analysis of 3,116,009 questions and 13,477,785 answers in Yahoo! Chiebukuro, a Q&A community, it was confirmed that (1) if questions that have a unique answer receive a right answer, then no additional answer tends to be submitted, (2) answers that have external references tend to be selected as the best answers, and (3) according to categories of questions, time zones when questions and best answers are submitted frequently differ. Based on the results, the following three could serve as trust indexes: (i) elapsed time from the submittal of answers or page view, (ii) existence of external references in an answer, (iii) the differences in submittal time of questions and answers in some question categories.

[キーワード]

Q&A コミュニティ、Yahoo!知恵袋、信頼性指標、情報爆発

1 はじめに

Web2.0 と呼ばれる言葉が象徴するように、ウェブ上での情報の流通形態が大きく変化してきている。以前も個人がホームページを開設するなどの方法で情報発信をおこなうことは可能であった。しかしながら、ツールの使い勝手の悪さやある程度の専門知識が必要であったこともあり、情報を自由に発信することができたのは限定された範囲の人々のみであったといえる。近年、Blog、SNS、Wiki などをはじめとする情報発信および情報共有手段の普及により、以前は固定化していた情報の送り手と受け手の境目が、現在ではますます曖昧になってきている。情報発信・共有にかかる労力や手続きが以前に比べ相当削減され、誰もが容易に情報を発信し不特定の第三者と情報を共有することができるようになったためである。

Blog、SNS、Wiki などの情報発信・共有技術は、CGM (Consumer Generated Media) とも呼ばれ、ウェブを活用して個人 (消費者) が情報 (コンテンツ) を生成していくための手段となっている。CGM は、かつて消費者が企業やマスメディアから提供される製品や情報を金銭的な対価を支払って消費するのみの存在であった時代との対比としてしばしば用いられる用語である。多くの「消費者」が「生産者」になるための手段を得た結果、多種多様かつ膨大な量の情報がウェブ上で日々発信されている。カリスマブロガーと呼ばれるような、専門家にも引けを取らない豊富な知識を有し人々に大きな影響を与える個人もすでに数多く登場してきている。さらには、個人の日記やメモ書きレベルの情報もウェブ上に大量に存在する状況となっている。

個人の情報発信・共有能力が飛躍的に向上し、玉石混濁の情報がウェブ上に拡散し続ける現在の状況は、「情報爆発」と呼ばれることもある。「情報爆発」が起きている状況では、あらゆる情報が存在し入手可能な状態にあるにもかかわらず、必要かつ十分な情報を膨大な情報の中から取捨選択して入手する必要が生じる。例えば、あるレストランのお薦めメニューを知りたいとする。ウェブを検索すればそのレストランを訪れたことのある人々が薦める「異なる」お薦めメニューが「大量に」存在する場合もある (どれが本当のお薦めメニューなのか判断し

づらい)。また、そのようなお薦めメニューは個人の主観的な意見が大きく反映されている場合もある（自分の嗜好に合ったものなのを半判断しづらい）。このように、個人の情報発信・共有能力の向上による恩恵（多様かつ大量の情報の入手）を得る一方で、弊害（大量の情報の中から適切な情報を取捨選択する必要性）も顕著に見られるようになってきており、入手する情報およびその情報の「信頼性」を確保することが今後極めて重要になると考えられている[5][9][11]。

本研究の最終的な目標は、情報爆発時代において個人が信頼性の高い情報を効率的に入手するための支援技術を構築することである。本論文ではその第一歩として、Q&A 型の情報共有コミュニティ（以降、Q&A コミュニティと呼ぶ）を対象とし、コミュニティ内で個人が入手する情報の信頼性を知るための手段（指標）を提供することを目的とする。続く2章では、Q&A コミュニティの関連研究について述べる。3章では、Q&A コミュニティにおける情報（回答）の信頼性評価指標構築へ向けた分析方法を述べる。4章では、代表的なQ&A コミュニティの一つであるYahoo!知恵袋コミュニティを対象としておこなった分析結果を示すとともに、得られた結果について考察し信頼性指標について議論する。最後にまとめについて述べ本論文を結ぶ。

2 関連研究

Q&A コミュニティの特徴の一つは、個人が抱えている疑問を解消するために質問と回答をユーザ同士が相互扶助的におこなうという点である。企業のヘルプデスクのように十分な知識を有する固定的な「回答者」が存在するわけではなく、回答者は次の質問者にもなりうる。

このようなユーザ同士による相互扶助的なQ&A コミュニティは、Web2.0以前にも存在していた。オープンソースのウェブサーバソフトウェアを開発しているApacheプロジェクトのユーザコミュニティなどが代表例として挙げられる。知識の豊富なユーザは相対的に少数しか存在しない場合が多く、一部のユーザが多数のユーザの「回答者」としての役割を果たす必要があったため、特定のユーザがコミュニティから離脱した場合、コミュニティが崩壊するといったことが多く見られた[6]。

そのため、豊富な知識を有するエキスパートユーザをコミュニティに長く引き留めておくための「動機付け」、初心者がエキスパートへ成長し新たな初心者をサポートできるまでコミュニティに継続的に参加させるための「動機付け」、多くのユーザを集めコミュニティを活性化させるための「動機付け」など、ユーザのコミュニティ参加に関する「動機付け」が研究の中心的対象の一つとなっている[2][4][8][10]。

現在多くのQ&A コミュニティにおいてユーザのコミュニティへの参加の動機付けのために採用されている手法には、ポイントやランキングがある。ユーザがおこなった回答を閲覧した他のユーザが評価しポイントを与える方法や、回答数の多さに応じてユーザをランキングし表彰する方法である。本論文で信頼性指標の妥当性評価のために用いるYahoo!知恵袋コミュニティにも同様の仕組みが用いられている。

Q&A コミュニティに参加するユーザは、ポイントやランキングを参考にして他のユーザの回答の質について推測することができる。しかしながら、前章で述べたように必ずしも自分にとって「信頼」できる有用な情報となっているかどうかを保証するものではない。この問題を解決するために、任意のユーザの間の信頼を評価するための手法[7]や質の高い情報を検索するための方法[1][3]についての研究が盛んにおこなわれている。本論文は、Q&A コミュニティにおける回答者および回答に対する信頼性を評価するための指標を提案しようとする点においては、先行研究と問題意識を共有していると言える。

Q&A コミュニティ、特にYahoo!知恵袋コミュニティにおけるユーザの参加理由や回答の質について調べた研究には[12][13][14][15]などがある。文献[12]は、Yahoo!知恵袋におけるQ&A コミュニティの利用者像を調べたものであり、Yahoo!知恵袋は活発に情報が交換・蓄積されているQ&A コミュニティであることを示している。文献[13]および[14]は、それぞれ質問タイプと回答データをもとに利用者の質問、回答内容の傾向を調べたものである。質問者は回答を素早く得ることを重要視して答えが一意に決まるような質問を行っていること、回答者は質問に接した時点で既に自身が持っている情報を回答内容の主体にしていること、などが報告されている。さらに文献[15]は、質問、回答の投稿数の違いによる利用者の特徴と形態素解析を用いて利用者の投稿内容を分析している。質問の投稿数が多い利用者は他の情報探索手段を利用することは少なくなる傾向があり、一次情報源としてのQ&A コミュニティの信頼性が高まっていることを示唆している。これらの結果は、本論文で目指している信頼性指標構築の動機ともなっている。

3 信頼性指標構築に向けた分析方法

3.1 対象 Q&A コミュニティ

本研究で対象とする Q&A コミュニティは Web ページ上で、投げかけた質問とそれに対する回答を誰もが閲覧できる場であり、以下を満たす。本論文では、質問分類、質問者、回答者を中心とし、質問と回答を閲覧する者を含めた集合をコミュニティと呼ぶ。

- ・ 閲覧に ID は必要ないが、質問、回答のためには ID を取得する必要がある。また、質問、回答内容には対応する質問者、回答者の ID が付与され、公開される。ID は他と重複しない限り自由に設定できる。
- ・ 質問内容に応じた質問分類が複数存在し、質問者は質問分類を一つ選択した上で質問を投稿する。
- ・ 質問は Web ブラウザを通じて自然言語で入力される。質問をみた閲覧者は自身で質問に答えることができると判断すれば、自然言語で回答を入力する。回答の正確性は保証されない。一つの質問に複数の回答が付与される場合もある。
- ・ 質問は投稿された時刻によって時系列に並べられ、どの閲覧者（回答者）にも同様の順番で提示される。
- ・ 質問から一定期間後に質問者が回答としてもっとも適していると判断した一つの回答を最良回答として選択する。最良回答が選ばれると他の回答者は回答できなくなる。同時に、選ばれた最良回答は最良回答であることが公開される。
- ・ 最良回答は回答者 ID に紐付けられ、回答者 ID ごとに実績として公開される。回答者がよい回答をするためのインセンティブとなる。

質問と回答の例（最良回答が選ばれた後の状態）を図 1 に示す。図 1 中の 1 行目は質問者 ID Alice の質問「昨夜のカレーが余ったのですが、何か使い道がありますかね？」である。これに回答できると自身で判断した回答者 ID Bob が回答「カレーうどんにします。」をしている。上記の例では、質問者は回答者 ID Carol の回答「カレーピラフにするといいですよ。余ったご飯も使えますし。」を最良回答としている。図の例のように、最良回答が一意に定まらない場合でも、質問者によって最良回答であると選択されたものが最良回答として公開される。

質問者 (ID Alice)	昨夜のカレーが余ったのですが、何か使い道がありますかね？
回答者1 (ID Bob)	カレーうどんにします。
回答者2 (ID Carol)	カレーピラフにするといいですよ。余ったご飯も使えますし。
回答者3 (ID Dave)	冷凍庫に保存しておく、食べたくなった時に食べられます。
回答者4 (ID Eve)	カレーパスタ、カレーコロッケなどなど色々使えます。

最良回答に選ばれた回答

図-1 質問と回答の例

3.2 分析方針

正しい回答を知らない質問者が、どれを信頼し正しい回答と信じればよいか、あるいは、全て信頼に値しない回答であるかどうかを回答内容だけから判断するのは難しい。自信をもって書かれていると判断される回答であっても回答者の勘違いにより意図せず誤った内容を含んでいる可能性もある。また、正解であるかどうかを確認する手段は必ずしも存在するとは限らない。

本研究で構築を目指す信頼性指標は、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ものである。質問者が正しい回答かどうかを判断する際に参考になる情報を増やすために、回答内容だけでなくそれ以外の付随情報をも手がかりとして利用していると考え、付随情報を表 1 のように整理した。表 1 の 1 行目はコミュニティの構成要素に対応する。各列はコミュニティの構成要素が発するメッセージの観点である。例えば、役割「質問者」から発せられるメッセージは「質問」であり、質問には「外部参照」として、文献の引用、第三者の発言、URL などによって表される外部参照が存在する場合がある、また、質問には発せられた「時刻」が存在する。他にも、質問者の過去の質問内容や回答内容、最良回答に選ばれた回数の実績も信頼性の手がかりとなる付随情報として考えられるが、本論文ではその第一段階としてメッセージに関与するものに限定することにした。役割で 3 行目の見込み回答者は、閲覧者のうち正しい回答ができると自身で判断したが、実際には回答を

表-1 信頼性の手がかりとなり得る付随情報

役割	メッセージ	メッセージの外部参照	メッセージの時刻
質問者	質問	URL など	質問日時(c)
回答者(既回答者)	回答	URL など(b)	回答日時
見込み回答者(未回答者)	閲覧(a)	なし	閲覧日時
閲覧者	閲覧	なし	閲覧日時

しなかった者を指している。ある回答者が回答しようとした回答と同一の回答が投稿されていたり、よりふさわしい回答が投稿されているとその回答者が判断したりした場合、回答を控える回答者を見込み回答者と呼ぶ。表1のうち、(a)、(b)、(c)に関して次のような仮定をおき分析した。

(a) 見込み回答者のメッセージ

既にふさわしい回答が存在する場合には見込み回答者は回答を避けると考えた。特に、質問に対しふさわしい回答が1件だけ投稿されている場合には、追加の回答が投稿されにくいと考える。図2(a)はその様子をあらわしており、ある回答者が質問にふさわしい回答(回答1)を投稿している場合に、それを見た見込み回答者は自身の回答(回答2)を投稿しない。ただし、単純に回答が少ないだけで他の見込み回答者が存在していない場合を考慮する必要がある。実際にこの傾向があるならば、一定時刻 t が経過している、あるいは、一定以上の閲覧数 v を超えているかどうか、最後の回答の信頼性の信頼性指標になり得ると考えた。

(b) 回答の外部参照

外部参照が含まれる回答は客観性があると判断され正しい回答として選ばれやすいと考えた。図2(b)はその様子を表している。回答1や回答3には外部参照が含まれておらず、回答2に外部参照が含まれている。質問者は外部参照先を第三者の意見として考えることができ、回答が正しいかどうかを判断するために利用できる場合があると考えた。実際にこの傾向があるならば、回答に含まれる外部参照の有無が信頼性指標になり得ると考えた。

(c) 質問のタイミング

Q&A コミュニティに寄せられる質問の数(コミュニティの規模)が大きくなると、回答者が全ての質問を閲覧することができなくなる場合がある。そこで、正しい回答ができる回答者が多く閲覧している時間帯が質問分類ごとに存在し、その時間帯に質問を投稿すれば正しい回答を得られる可能性が高いと考えた。図2(c)はその様子を表しており、正しい回答が得られる時間帯に質問に正しい回答が得られる可能性が高いことを示している。実際にこの傾向があるならば、質問分類、質問時刻、回答時刻の組合せが信頼性指標になり得ると考えた。

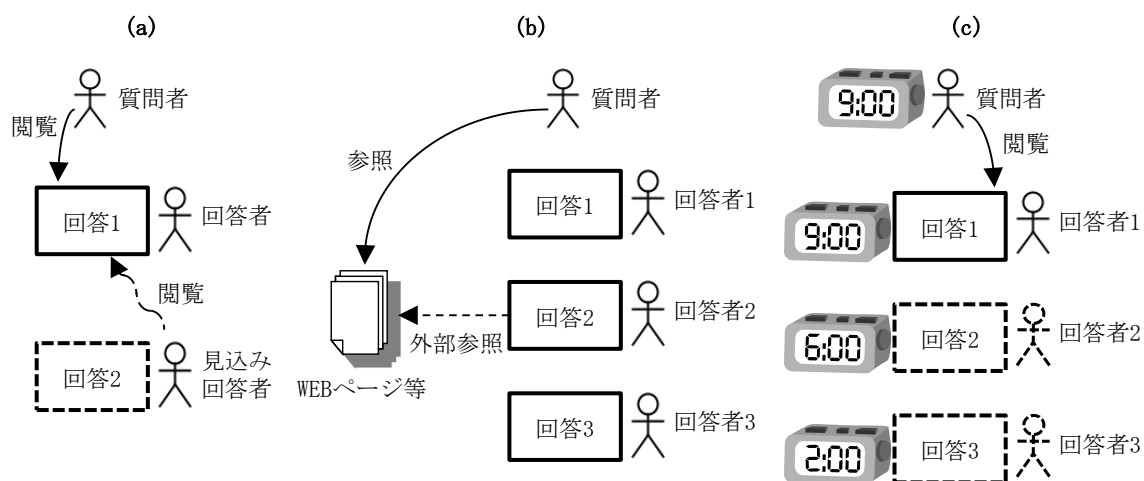


図-2 最良回答の選択に関する仮説の例

4 信頼性指標構築に向けた評価

4.1 対象 Q&A コミュニティ

対象 Q&A コミュニティは Yahoo!知恵袋とした。評価対象データは、国立情報学研究所から提供された知識検索サービス「Yahoo!知恵袋」で実際に投稿された質問、回答である。2004 年 4 月 1 日～2005 年 10 月 31 日に投稿された利用者 242,334 人の質問と回答から構成されており、質問数 3,116,009 件、回答数 13,477,785 件である。各々の質問、質問者、回答、回答者には ID が付与されている。また、各質問にはカテゴリ番号、カテゴリ名、質問文、質問投稿日、質問最終更新日が含まれ、各回答には回答投稿日、回答最終更新日、回答文、参考 URL (最大 3 件) が含まれる。カテゴリとは、質問内容による質問分類であり対象データでは 286 カテゴリが定義されている。例えば、「政治、社会問題」「自動車」のような質問分類がある。

4.2 評価方法

4.2.1 (a) 見込み回答者のメッセージ

本評価では、決着回答が質問に対して適切な回答であったかどうかを実際の質問者、回答者とは異なる被験者により調べる。ここで、決着回答を次のとおり定義する。ある質問 q_i に対する回答の集合 $A_i\{a_1, a_2, \dots, a_{n_i}\}$ (n_i は q_i への回答の総数をあらわす) のうち、最後に投稿された回答 a_{n_i} が最良回答として選ばれている場合に a_{n_i} を決着回答と呼ぶ。なお、集合 A_i の添え字は回答の投稿日時によって定まるが、今回の対象データでは、投稿日時のタイムスタンプにばらつきがあったため、 $n_i = 1$ のものを選び、評価の対象とした。つまり、ある質問に 1 件だけの回答が投稿され、それが最良回答として選ばれているものに限定した。ただし、回答が 1 件だけの場合に、見込み回答者が回答を避けたのか、単に見込み回答者が現れなかったのかどうかを区別することが難しいため、被験者を用いた評価を実施した。

実験には、100 件の $n_i = 1$ の決着回答と質問 q_i の対を用いた。 $n_i = 1$ の決着回答には、見込み回答者が a_i の回答内容を正しい回答と評価したのではなく、ただ単に見込み回答者が現れなかった可能性がある。そのため、100 件のうち、50 件は質問、回答の総件数が多い質問分類から、残りの 50 件は質問、回答の総件数が小さい質問分類から無作為に選んだ。具体的には、質問、回答の総件数上位 10 質問分類 (161319 件の $n_i = 1$ の決着回答と質問) から 50 件、質問、回答の総件数が下位 (1000 件未満) の 158 件の質問分類の質問・決着回答 (885 件の $n_i = 1$ の決着回答と質問) から無作為に 50 件を選んだ。具体例としては、「ドラマ『鏡は眠らない』で、三田佳子の義理の娘の役をしていた女優さんの名前って何でしたっけ?」という質問に対して付けられた「民子のことですよ? 民子なら野波麻帆さんですよ。」という決着回答である。

被験者は奈良先端科学技術大学院大学情報科学研究科博士前期課程の学生 5 名である。また、被験者には次に示すような質問に答えてもらった。

1. 被験者への質問: 質問に対する正解を知っているか? 回答の選択肢: 知っている、知らない
2. 被験者への質問: 一意に答えが決まる質問であるか? 回答の選択肢: 一意に決まる、一意に決まらない
3. 被験者への質問: 回答がベストアンサーとしてふさわしいか? 回答の選択肢: ふさわしい、ふさわしくない

4.2.2 (b) 回答の外部参照

外部参照の有無を URL 存在の有無と定義し評価した。対象データには、1 件の回答につき 3 件までの URL が付記されているため、これを利用した。URL を含む回答の最良回答の件数と含まない回答の最良回答の件数を求める。

4.2.3 (c) 質問のタイミング

質問分類ごとに、質問時間帯のピークと回答時間帯のピークを調べ、その差を算出する。ただし、対象データ全体に対する比率とする。具体的には、質問が投稿された時刻 (0~23 時) をもとに、1 時間おきの時間帯ごとに質問の集合 Q_{ct} を集計する。 c は質問分類であり、今回の対象データでは 286、 t は時間帯であり、 $0 \leq t \leq 23$ である。同様に回答の集合 A_{ct} も求める。また、ある時間帯 t の全体での質問数 $\sum |Q_{it}|$ ($i \in C$: C は全ての質問分類の集合) に対するある質問分類での質問数 $|Q_{ct}|$ の比率を求め、その値がもっとも大きくなる t_{cq} を質問が最大となる時間帯とする。同様にして、回答が最大となる時間帯 t_{ca} を求める。質問分類によって $|t_{cq} - t_{ca}|$ の大きさが変わるかどうかを求める。評価は時間帯の質問、回答の総数の最大値が 1000 以上である質問分類を対象とした。

表-2 質問、回答の総件数が上位の質問分類に含まれる決着回答の評価

正解を知っている	一意に答えが決まる	最良回答としてふさわしい	回答数
YES	YES	YES	28(11.2%)
		NO	5(2.0%)
	NO	YES	9(3.6%)
		NO	6(2.4%)
NO	YES	YES	83(33.2%)
		NO	32(12.8%)
	NO	YES	37(14.8%)
		NO	50(20.0%)

表-3 質問、回答の総件数が下位の質問分類に含まれる決着回答の評価

正解を知っている	一意に答えが決まる	最良回答としてふさわしい	回答数
YES	YES	YES	45(18.0%)
		NO	8(3.2%)
	NO	YES	6(2.4%)
		NO	8(3.2%)
NO	YES	YES	104(41.6%)
		NO	27(10.8%)
	NO	YES	35(14.0%)
		NO	17(6.8%)

表-4 回答の URL 含有率

URL 含有数	最良回答	最良回答以外
含まない	2731581(87.66%)	10037932(96.87%)
一つ含む	363354(11.66%)	307023(2.96%)
二つ含む	16837(0.54%)	13621(0.13%)
三つ含む	4236(0.14%)	3201(0.03%)

4.3 結果

4.3.1 (a) 見込み回答者のメッセージ

結果を表 2、表 3 に示す。表 2 は質問、回答の総件数が上位の質問分類に含まれる決着回答、表 3 は質問、回答の総件数が下位の質問分類に含まれる決着回答の結果である。質問分類に属する質問、回答の総件数による大きな違いは見られなかった。被験者が正解を知っている場合で、一意に答えが決まるものについては、質問分類に限らず、決着回答が信頼できる回答である可能性が高いといえる。質問、回答の総件数が上位のもので、一意に答えが決まるもののうち、ベストアンサーとしてふさわしいと被験者に判断されたものは 28 件であるのに対して、ふさわしくないと判断されたものは 5 件、また、質問、回答の総件数が下位のもので、最良回答としてふさわしいと判断されたものが 45 件、ふさわしくないと判断されたものは 8 件であった。一方、一意に答えが定まらないものについては、決着回答が信頼できるとは限らない結果が得られた。

4.3.2 (b) 回答の外部参照

回答全体から URL を含む回答を調べたところ、表 4 に示される結果が得られた。回答が最良回答として選ばれたかどうかに関わらず、多くの回答は URL による外部参照を含まないことがわかった。ただし、最良回答に選ばれた回答のうち、URL を 1 件以上含むものの比率は、12.34%であったのに対し、最良回答以外では、URL を 1 件以上含むものの比率は 3.12%であった。

4.3.3 (c) 質問のタイミング

結果を図 3 に示す。図 3 では、横軸は質問のピーク時間帯と最良回答のピーク時間帯の差（時刻差）を、縦軸は各時刻差に対する質問分類の数を表している。35 件の質問分類において質問のピーク時間帯と最良回答のピー

ク時間帯は同一であった。質問のピーク時間帯と最良回答のピーク時間帯の差が4時間以上ある質問分類は29件であった。時刻差が大きい質問分類として、「楽器」「ゲーム」があった。これら2件の時間帯と質問、最良回答件数の全体に占める割合を図4に示す。図4(a)は「楽器」、図4(b)は「ゲーム」である。図4、図5では、横軸は時間帯を、縦軸は対象データ全体の同一時間帯の質問件数、及び、最良回答件数に占める各質問分類の質問件数、最良回答件数の割合を表している。図4(a)では、質問の比率が最大の時間帯が22時、最良回答の比率が最大の時間帯が4時であり、その差は6時間である。また、図4(b)では質問の比率が最大となる時間帯は18時、最良回答の比率が最大となる時間帯は5時であり、時刻差は11時間であった。また、時間帯ごとの質問の比率、最良回答の比率の推移も異なる。一方、図5に示した「恋愛相談、人間関係の悩み」「料理、グルメ、レシピ」の質問分類では、質問、最良回答の比率が最大の時間帯が同じであり、質問の比率と最良回答の比率の推移も類似している。

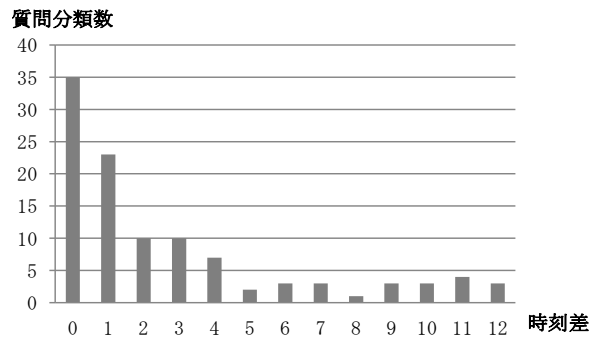


図-3 質問、最良回答のピーク時間帯の差

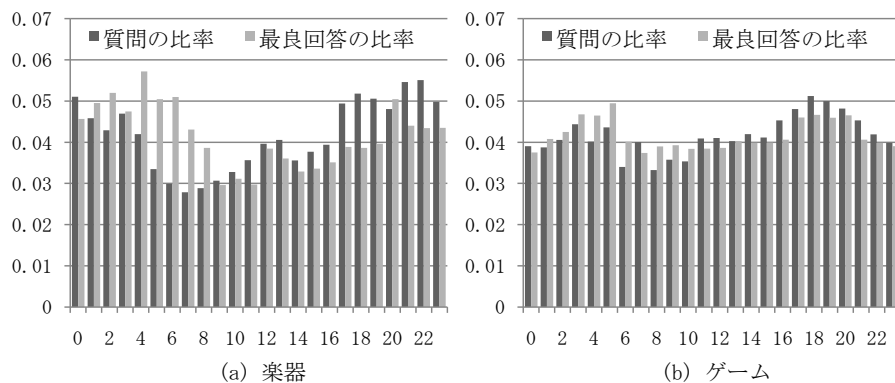


図-4 質問のピーク時間帯と最良回答のピーク時間帯の差が大きい質問分類の例

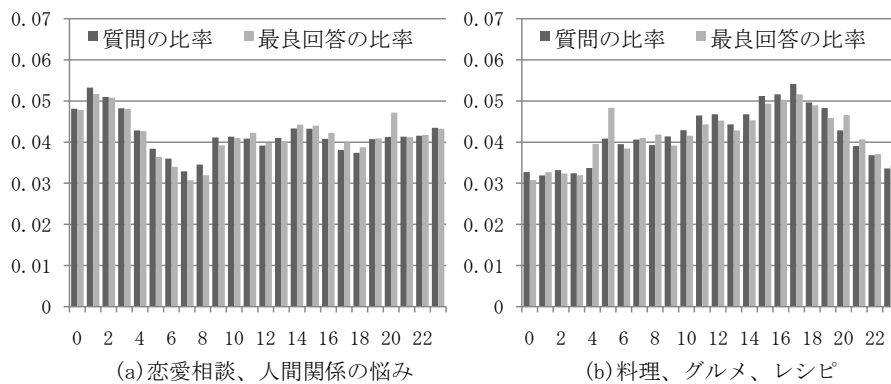


図-5 質問のピーク時間帯と最良回答のピーク時間帯の差が小さい質問分類の例

4.4 考察

本評価により、回答の信頼性指標につながる手がかりを得ることができた。以下に3件それぞれについて信頼性指標への展望を与えつつ、考察する。

(a) 見込み回答者のメッセージ

既に質問にふさわしい回答が投稿されているときには、見込み回答者が回答を追加しない傾向があることを被験者の実験によりうかがうことができた。特に、回答が一意に定まる場合にその傾向が強まる評価結果が得られた。回答が一意に決まる質問をしているかどうかを事前に質問者に申告してもらったり、その注意書きを含めたりすることにより、次のような信頼性指標で質問者が回答の信頼性を評価する支援ができる。回答が投稿されてから一定時刻 t が経過している、あるいは、一定以上の閲覧数 v を超えている。

(b) 回答の外部参照

最良回答とそれ以外の回答の間で、URL による外部参照が1件以上含まれる割合に4倍近い差があった。外部参照は質問者が最良回答を選ぶ際の手がかりの一つとなっているといえる。今回の評価は計算機で扱いやすいよう URL のみを対象としたが、目視による調査によって次のような外部参照も存在した。図6はURL以外の外部情報を含む回答の例であり、1行目は質問者による質問内容を、2行目から5行目はその質問に対する回答者1から回答者4までのそれぞれの回答を表している。図6に示すような英語に関する質問に対する回答において、「隣の外国人が言っています」という外部参照が含まれている回答者1の回答が最良回答に選ばれていた。なお、いずれの回答も正解ではあるが、回答者1の回答が質問者により最良回答と選択された。何らかの方法でURL以外の外部参照をも計測できれば、外部参照の有無は有力な信頼性指標になり得ると考えられる。また、回答者にはなるべく外部参照を使いながら回答をすることを促すことで回答の客観性を示すことができる等のアドバイスができると期待される。

(c) 質問のタイミング

質問分類によっては、質問が投稿される時間帯のピークと最良回答が投稿される時間帯のピークが異なる結果が得られた。図4に示した「楽器」と「ゲーム」の質問分類においては、回答の時間帯が信頼性指標となり得る。また、質問の投稿時間帯を意図的にずらすことにより、より適切な見込み回答者に質問を閲覧される可能性が高まる。一方、質問が投稿される時間帯のピークと最良回答が投稿される時間帯のピークが同一の質問分類もあることがわかった。図5に示したように「恋愛相談、人間関係の悩み」「料理、グルメ、レシピ」のほかにも「子育て、出産」「妊娠、出産」といった既婚の女性が主な質問者・回答者であると考えられる質問分類でも同様の傾向があった。

5 まとめ

Q&A コミュニティを対象とし、情報（回答）の信頼性を評価するための信頼性指標構築に向けた分析を実施した。(1)回答が一意に定まるような質問に対し正しい回答が投稿されている場合には、追加の回答が投稿されにくい傾向があること、(2)URLによる外部参照を含む回答はそうでない回答と比較して最良回答に選ばれる傾向があること、(3)質問分類によって多くの質問が投稿される時間帯と最良回答が投稿される時間帯が一致するもの、乖離があるものがあることをYahoo!知恵袋のデータを用いて確認した。また、その結果にもとづき、(i)回答が一

質問者	英語で「三分之一」ってなんて言うのでしょうか？
回答者1	隣の外国人に聞いたら、one third と教えてくれました。
回答者2	one third です。日本語は3分の1と書きますが、英語の場合は逆になります。
回答者3	a third、または、one third です。
回答者4	one third です。序数が分母、基数が分子として分数が表せます。

最良回答に選ばれた回答

図-6 URL以外の外部情報を含む回答の例

意に定まるタイプの質問では、最後の回答が投稿されてからの経過時間やその回答の非閲覧数が信頼性指標になり得る、(ii)回答を裏づけるような外部参照の有無は信頼性指標になり得る、(iii)多くの質問が投稿される時間帯と多くの最良回答が投稿される時間帯が異なる質問分類では、質問や回答の時間帯が信頼性指標になり得ることが推測される。

謝辞

本論文でおこなった実験にご協力頂いた奈良先端科学技術大学院大学情報科学研究科の皆様へ深く感謝する。なお、本論文では、Yahoo!株式会社が国立情報学研究所に提供したYahoo!知恵袋データセットを使用した。本研究の一部は、文部科学省「次世代IT基盤構築のための研究開発」の委託に基づいて行われた。また、本研究の一部は、文部科学省科学研究補助費（若手B：課題番号20700028）による助成を受けた。

[参考文献]

- [1] Agichtein, E., Castillo, C., Donato, D., Gionis, A. and Mishne, G.: Finding high quality content in social media, with an application to community-based question answering, Web Search and Data Mining (WSDM), ACM Press, Stanford, USA, pp.183-194 (2008).
- [2] Beenen, G, Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P. and Kraut, R.E.: Using social psychology to motivate contributions to online communities, Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW'04), pp.212-221 (2004).
- [3] Bian, J., Liu, Y., Agichtein, E. and Zha, H.: Finding the right facts in the crowd: factoid question answering over social media, Proceeding of the 17th International Conference on World Wide Web (WWW'08), pp.467-476 (2008).
- [4] Danis, C.M., Viegas, F.B., Wattenberg, M. and Kriss, J.: Your place or mine?: visualization as a community component, Proceeding of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'08), pp.275-284 (2008).
- [5] Grabner-Kräuter, S., Kaluscha, E.A. and Fladnitzer, M.: Perspectives of online trust and similar constructs: a conceptual clarification, Proceedings of the 8th International Conference on Electronic Commerce (ICEC'06), pp.235-243 (2006).
- [6] Lakhani, K.R. and Hippel, E.: How open source software works: “free” user-to-user assistance, Research Policy, Vol.32, No.6, pp.923-943 (2003).
- [7] Liu, H., Lim, E., Lauw, H.W., Le, M., Sun, A., Srivastava, J. and Kim, Y.A.: Predicting trusts among users of online communities: an epinions case study, Proceedings of the 9th ACM Conference on Electronic Commerce (EC'08), pp.310-319 (2008).
- [8] Nonnecke, B., Preece, J. and Andrews, D.: What lurkers and posters think of each other, Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04), pp.70195a (2004).
- [9] Salam, A.F., Iyer, L., Palvia, P. and Singh, R.: Trust in e-commerce, Communications of ACM, Vol.48, No.2, pp.72-77 (2005).
- [10] Tedjamulia, S.J.J., Dean, D.L., Olsen, D.R. and Albrecht, C.C.: Motivating content contributions to online communities: toward a more comprehensive theory, Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05), pp.193b (2005).
- [11] Uslaner, E.M.: Trust online, trust offline, Communications of ACM, Vol.47, No.4, pp.28-29 (2004).
- [12] 三浦麻子、川浦康至、地福節子、大瀧直子、岡本真：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回論文集、3D3-1 (2006)。
- [13] 三浦麻子、川浦康至、地福節子、大瀧直子、岡本真：知識共有コミュニティを創り出す人たち(2)：「質問タイプ」から見た参加行動、日本社会心理学会47回大会論文集、pp.496-497 (2006)。
- [14] 三浦麻子、川浦康至、地福節子、大瀧直子、岡本真：知識共有コミュニティを創り出す人たち(3)：「回答者」データから見るコミュニティ内の「知識」、日本社会心理学会47回大会論文集、pp.498-499 (2006)。
- [15] 三浦麻子、川浦康至、地福節子、大瀧直子、岡本真：知識共有コミュニティを創り出す人たち(4)：形態素解析による「知識」の分析、人工知能学会全国大会第21回論文集、3D3-1 (2007)。