

QA コミュニティにおける複数情報源を用いた効果的な質問推薦

甲谷 優[†] 岩田 具治^{††} 塩原 寿子[†] 藤村 考[†]

[†] 日本電信電話株式会社 NTT サイバーソリューション研究所

〒 239-0847 神奈川県横須賀市光の丘 1-1

^{††} 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

〒 619-0237 京都府相楽郡精華町光台 2-4

E-mail: [†]{kabutoya.yutaka,shiohara.hisako,fujimura.ko}@lab.ntt.co.jp, ^{††}iwata@cslab.kecl.ntt.co.jp

あらまし QA コミュニティの回答者の知識, 興味, 嗜好を推定し, 質問を推薦する手法を提案する. 効果的な質問推薦ができれば良質な回答が増え, その結果 QA コミュニティ全体の活性化につながると我々は考える. ここで, 効果的な質問推薦を, (1) カテゴリよりも適切な質問を提示でき, (2) 回答がついていない質問であっても推薦でき, (3) 回答したことのないユーザにも推薦できるものであると考えた. このような効果的な質問推薦を実現するために, 我々は複数の推薦アルゴリズム, 複数の情報源によりユーザの知識, 興味, 嗜好を推定し, それらをロジスティック回帰モデルにて統合する. 本稿では教えて! goo の実ログデータを使って, 提案手法により効果的な質問推薦が実現できることを示す.

キーワード 質問推薦, QA コミュニティ, ロジスティック回帰モデル

Effective Question Recommendation Using Multiple Features for Question Answering Communities

Yutaka KABUTOYA[†], Tomoharu IWATA^{††}, Hisako SHIOHARA[†], and Ko FUJIMURA[†]

[†] NTT Cyber Solutions Laboratories, NTT Corporation

1-1 Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847 Japan

^{††} NTT Communication Science Laboratories, NTT Corporation

2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0237 Japan

E-mail: [†]{kabutoya.yutaka,shiohara.hisako,fujimura.ko}@lab.ntt.co.jp, ^{††}iwata@cslab.kecl.ntt.co.jp

Abstract We propose a new method of recommending questions to answerers so as to suit the answerers' knowledge, interests, and preference in User-Interactive Question Answering (QA) communities. A question recommender can help activation of a QA community. An effective question recommender should satisfy the following three requirements: First, its accuracy should be higher than the existing category-based approach. Second, it should be able to recommend unanswered questions. Third, it should be able to support even those people who have never answered a question previously. To achieve an effective question recommender, we use combining QA users' knowledge, interests, and preference by a logistic regression model. Experiments on real log data sets of a famous Japanese QA community, Oshiete goo, show that our recommender satisfies the three requirements.

Key words Question Recommendation, Question Answering Communities, Logistic Regression Model

1. はじめに

近年, 教えて! goo^(注1) 等の QA コミュニティでは, 毎日 2,000 件以上もの質問が投稿されており, 注目を集めている.

質問が投稿されると, 親切な他者により回答が付与され, 質問者の悩みや疑問が解決する. さらに, その回答は質問者だけでなく, 同じ疑問や悩みを抱えた別のユーザにも有益な情報で再利用されることが多い. このように, 多くの回答者の集合知を多くのユーザに共有できる, 有用性の高いサービスである.

QA コミュニティにおいて, 常に回答に満足している質問者

(注1) : <http://oshiete.goo.ne.jp/>

ばかりではないと我々は考え、そのような回答に満足していない質問者がどれくらい存在するのか、アンケート調査を行った。結果、QA コミュニティにおける質問者の多くは、常に回答に満足しているわけではないことがわかった。

このような質問者の不満を解消するために、質問推薦を提案する。適切な質問が推薦され、ユーザがより早くそのような質問にたどり着くことができれば、回答のモチベーションは上がり、回答数が増えるだろう。すなわち、我々は質問推薦は QA コミュニティ全体の活性化に有効であると考え。また、回答数の増加は、良質な回答の増加にもつながっていくため、質問推薦は回答の質の向上という効果も期待できる。

我々は、回答するのに適切な質問とはどのような質問なのかについてもアンケート調査を行った。そのアンケート調査の結果から、我々は回答するのに適切な質問とは、ユーザの知識、興味、嗜好に則した質問であると結論付けた。ここで、ユーザの知識、興味、嗜好をそれぞれユーザにとって既知の話題、ユーザにとって知りたい話題、話題とは関係のないユーザの回答基準と定義する。

一方で、現在の QA コミュニティでは 50% 以上の回答者が質問に回答する際にカテゴリ^(注2)を用いていることがわかった。しかしながら、カテゴリでは回答者が自身の知識、興味、嗜好に合った質問を発見するのは困難である。理由は 2 点ある: 1 つ目は、カテゴリ数が膨大で、また互いに類似するカテゴリも数多く存在することである。そのため、回答者に適したカテゴリを選択することが困難である。2 つ目は、カテゴリ内の質問数もやはり膨大であることである。したがって、カテゴリ内の膨大な質問から、回答するのに適した質問を選択するのも困難である。たとえば、「淡路島の釣りスポット」に関する質問は、「関西」カテゴリに投稿すべきであろうか、それとも「釣り」カテゴリに投稿すべきであろうか。ここでもし前者のカテゴリに投稿された場合、後者のカテゴリをよく利用する適切な回答者に質問が回答されないことになる。すなわち、多くの潜在的な回答が失われている可能性がある我々は考える。

したがって、QA コミュニティにはカテゴリよりもユーザの知識、興味、嗜好にあった質問を提示する機能が必要であると考え。そのような質問推薦を実現するために、本研究では複数の情報源を統合した手法を提案する。まず、ユーザの知識の推定に、ユーザの回答した質問 (以後、回答履歴と呼ぶ) に基づくコンテンツフィルタリングを行う。これは、ユーザが回答した質問の内容は既知の話題であると考えたためである。同様に、ユーザの興味の推定にはユーザ自身の質問 (以後、質問履歴と呼ぶ) に基づくコンテンツフィルタリングを用いる。これは、ユーザは自身の知りたい話題について質問すると考えたためである。これは、ユーザの嗜好の推定には協調フィルタリングを用いる。これは、話題とは関係のないユーザの回答基準が、回答履歴の類似した別のユーザから推定できると考えたためである。

(注2): 各質問はそれぞれ単一のカテゴリに属している。たとえば、「タマネギの調理のしかた」に関する質問は「料理」カテゴリに属する

表 1 調査対象と質問数・回答数の分布

質問数	回答数				
	0	1-4	5-10	11- 回	計
0	216	418	88	104	826
1-4	523	0	0	0	523
5-10	71	77	143	93	384
11-	21	24	53	230	328
計	831	519	284	427	2061

本稿の構成は以下の通りである。まず、第 2 章にて QA コミュニティの利用状況に関するアンケート調査から質問者の不満とその要因、回答者の動機、回答者がどのようにして質問にたどり着くかを明らかにする。次に、第 3 章にて質問推薦の満たすべき要件を明らかにし、その上での提案手法について述べる。第 4 章では、教えて!goo の実ログデータを用いた実験を通して、提案手法が既存の推薦手法やカテゴリよりも高精度であることを示す。第 5 章では、さらに分析実験を行い、提案手法が質問推薦の要件を満たすことを示す。第 6 章では、関連研究について言及し、本研究の位置付けを行う。第 7 章では、まとめと今後の課題について言及する。

2. QA コミュニティの利用状況に関する調査

2.1 アンケート調査の対象

我々は、QA コミュニティの利用状況に関して、goo リサーチ^(注3)を利用してアンケート調査を行った。表 1 に、調査対象の QA コミュニティにおける質問数と回答数の分布を示す。表の通り、今回は質問数と回答数の多いユーザも調査対象とすることができた。

2.2 質問者の満足度

まず、質問についた回答に対する満足度に関する問いを設けた。対象は QA コミュニティで 1 度でも質問したことのある 1235 名とした。項目は、「毎回満足」、「ほぼ満足、たまに不満」、「不満なときもある」、「ほぼ不満、たまに不満」、「毎回不満」の 5 つを設けた。回答形式は単一選択式とした。

本設問の回答から、QA コミュニティにおいて 1 度でも質問したことのある 1235 人のユーザの中で、質問についた回答に「毎回満足」と答えたユーザは 103 人と 8% 足らずであった。逆に、90% 以上のユーザがついた回答に不満なときもあるということがわかった。

さらに、質問についた回答に不満を抱く要因に関する問いを設けた。対象は、ついた回答に不満なこともあった質問者 1132 名とした。回答形式は複数選択式とした。

結果を、図 1 に示す。約 50% が「期待したとおりの回答ではなかったから」、「ついた回答では問題が解決しなかったから」と、ついた回答に関する質に不満を抱いていた。また、約 25% が「十分な数の回答が得られなかったから」と、ついた回答の量に関して不満を抱いていた。また、10% 程度ではあるが「回答が遅かったから」と、回答がつくまでにかかる時間の遅さを不満の理由に挙げるユーザも存在した。

(注3): <http://research.goo.ne.jp/>

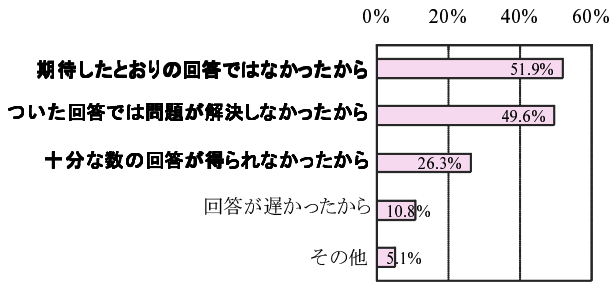


図 1 質問者の不満の要因 (計: 1132 名)

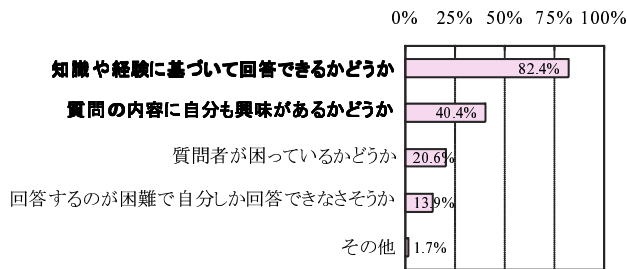


図 2 回答者が回答する動機 (計: 1230 名)

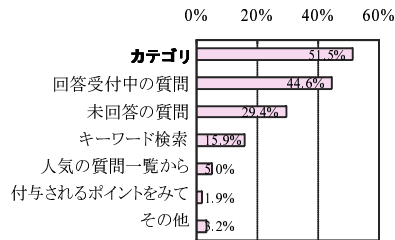


図 3 回答する質問を選択する手段 (計: 1230 名)

2.3 回答者が回答する動機

まず, QA コミュニティにおける回答者の, 質問に回答する動機に関する問いを設けた. 対象は, QA コミュニティで 1 度でも回答したことのある 1230 名とした. 回答形式は複数選択形式とした.

結果を図 2 に示す. 80% を越えるユーザが「知識や経験に基づいて回答できるかどうか」を選択し, 50% 近くのユーザが「質問の内容に自分も興味があるかどうか」を選択した. このように, 多くのユーザは自身の知識, 興味に合った質問に回答していることがわかる.

2.4 回答者が質問を選択する手段

さらに, 回答者が現在どのような基準で回答する質問を選択しているのかに関する問いを設けた. 対象は上記設問と同じく 1 度でも回答したことのある 1230 名とした. 回答方式は複数選択式とした.

結果を図 3 に示す. 約 50% のユーザが「カテゴリ」から選択すると答えた.

しかしながら, カテゴリによって回答する質問を選択するのは困難なときもあると考える. それは, 互いに類似したカテゴ

表 2 提案手法の説明に用いる変数

変数	説明
Q	質問集合
W	単語集合
q	質問
$\mathbf{q} = \{q_w\}_{w=1}^{ W }$	質問 q の単語特徴ベクトル
U	ユーザ集合
U_q	質問 q に回答したユーザ集合
u	ユーザ
$\theta_u^A = \{\theta_{uw}^A\}_{w=1}^{ W }$	ユーザ u の質問した質問における単語の特徴量
$\theta_u^R = \{\theta_{uw}^R\}_{w=1}^{ W }$	ユーザ u の回答した質問における単語の特徴量
Q_u^A	ユーザ u が質問した質問集合 (質問履歴)
Q_u^R	ユーザ u が回答した質問集合 (回答履歴)
r	ユーザが質問に回答したか
a	ユーザが質問したか
c_q	質問 q が属するカテゴリ
f	素性
F	素性集合

りも数多く存在し, かつ特定のカテゴリではそこに属する質問の数が膨大であるためである. したがって, 回答者が質問を選択する手段として, カテゴリは不十分であると考えらる.

3. 提案手法

ここで, 我々は QA ユーザを回答する可能性の高い質問へナビゲートするために, **質問推薦**を提案する. 質問推薦のシステムは各ユーザに対してそのユーザに対する適合度順に並んだ n 件の質問を提示する. ただし, そのユーザが既に回答した質問や, 回答が締め切られた質問は除く.

有効な質問推薦は, 以下に示す 3 つの要件を満たす必要がある.

- (1) 既存のカテゴリよりもユーザに適した質問を提示しなければならない. 2.4 節のアンケート調査の結果が示す通り, 50% 以上のユーザがカテゴリを用いて質問を選択している.
- (2) 回答したことのないユーザであっても推薦できることが望ましい. QA コミュニティの多くのユーザは, 回答したことがない. たとえば教えて!goo の場合, 2009 年 6 月時点において 1 度でも回答したことのあるユーザが 253,072 名であったのに対し, 1 度も回答したことがない (ただし, 質問はしたことがある) ユーザは 327,669 名であった.
- (3) 未回答の質問であっても推薦できなくてはならない. QA コミュニティには毎日何千件もの新規の質問が投稿されるが, そのような新規の質問には回答がついていない. たとえば教えて!goo の場合, 1 日に 2,000 件以上もの質問が投稿される.

これらの要件を満たす質問推薦を実現するにあたり, 2 章のアンケート調査を踏まえ, 「ユーザは, 自身の**知識, 興味, 嗜好**に合った質問に回答する可能性が高い」と仮説を立てる.

今, 表 2 に本節で用いる各変数とその説明について記す. ただし, 単語の特徴量とは, **tf-idf** を指す. idf のドキュメント

集合には、質問集合 Q を用いる。

3.1 ユーザ知識の推定

本研究では、ユーザの知識を推定するために、回答した質問(回答履歴)の内容に関する情報を用いて、Mooney ら [1] により提案されたコンテンツフィルタリングを行う。質問の内容情報としては、以下に示すように本文とカテゴリを用いる。

(1) 回答履歴の本文を用いた知識推定

質問の本文の bag-of-words に含まれる各語について、特徴量 (tf-idf 値) を算出し特徴ベクトルとする。ユーザの回答履歴との特徴ベクトルのコサイン類似度により、各質問がユーザの知識にどれほど合致しているかを算出する。

$$P_{\text{ContentA}}(r = 1|u, q) = \cos(\mathbf{q}, \boldsymbol{\theta}_u^R). \quad (1)$$

本手法により、回答がついていない質問であっても推薦できる (要件 3)。

(2) 回答履歴のカテゴリを用いた知識推定

ユーザの回答履歴から、カテゴリの回答頻度を算出する。回答頻度の高いカテゴリに属する質問をユーザの知識に合致した質問とみなす。

$$P_{\text{CategoryA}}(r = 1|u, q) = P(r = 1|u, c_q). \quad (2)$$

本手法により、少なくともカテゴリ程度には高精度な推薦ができ (要件 1)、また回答がついていない質問であっても推薦できる (要件 3)。

3.2 ユーザ興味の推定

本研究では、ユーザの興味を推定するために、ユーザ自身の質問(質問履歴)の内容に関する情報を用いてコンテンツフィルタリングを行う。質問の内容情報としては、知識の推定と同様に本文とカテゴリを用いる。

(3) 質問履歴の本文を用いた興味推定

ユーザの質問履歴との特徴ベクトルのコサイン類似度により、各質問がユーザの興味にどれほど合致しているかを算出する。

$$P_{\text{ContentQ}}(r = 1|u, q) = \cos(\mathbf{q}, \boldsymbol{\theta}_u^A). \quad (3)$$

本手法により、回答したことのないユーザであっても推薦でき (要件 2)、回答のついていない質問でも推薦できる (要件 3)。

(4) 質問履歴のカテゴリを用いた興味推定

質問頻度の高いカテゴリに属する質問をユーザの興味に合致した質問とみなす。

$$P_{\text{CategoryQ}}(r = 1|u, q) = P(a = 1|u, c_q). \quad (4)$$

本手法も質問履歴の本文を用いた手法と同様、要件 2 と要件 3 を満たす。

3.3 ユーザ嗜好の推定

ユーザの嗜好とは、回答する質問を選択する基準になる、質問の内容以外に関する情報と定義する。本研究では、そのようなユーザの嗜好を推定するのに、協調フィルタリングの手法を用いる。

(5) ユーザベースの協調フィルタリング

Resnick ら [2] によって提案された GroupLens というシステムに利用されている手法。似た回答履歴をもつユーザの嗜好から、ユーザの嗜好が推定される。

$$P_{\text{UserCF}}(r = 1|u, q) = \frac{1}{|U_q|} \sum_{u' \in U_q} \cos(Q_u^R, Q_{u'}^R). \quad (5)$$

本手法ではカテゴリやテキスト等の質問の内容情報は考慮しない。したがってカテゴリよりも高精度な推薦が可能になる可能性がある (要件 1)。

(6) アイテムベースの協調フィルタリング

Sarwar ら [3] によって提案された推薦手法。ユーザベースのものに比べ、計算コストが少なく済む。そのため、Amazon^(注4) 等、商用サービスで用いられている。

$$P_{\text{ItemCF}}(r = 1|u, q) = \frac{1}{|Q_u^R|} \sum_{q' \in Q_u^R} \cos(U_q, U_{q'}). \quad (6)$$

本手法を用いた理由は、ユーザベースの協調フィルタリングと同様で、カテゴリを越えた高精度な推薦が可能になる (要件 1) と考えたためである。

3.4 ロジスティック回帰モデル

本研究では、複数手法をロジスティック回帰モデルを用いて統合する。ロジスティック回帰モデルを採用した理由としては、Jin ら [4] の実験により推薦における優位性が確認されており、高精度な質問推薦を実現することが可能である (要件 1) と考えたためである。ロジスティック回帰モデルを質問推薦に適用した場合、回答者 u が質問 q に回答する確率は以下のように表すことができる。

$$P(r = 1|u, q) = \frac{1}{1 + \exp(-\sum_{f \in F} \lambda_f P_f(r = 1|u, q) - \lambda_0)}. \quad (7)$$

ただし、 $\boldsymbol{\lambda} = \{\lambda_f\}_{f \in F}$ は未知の重みパラメータである。

未知パラメータ集合 $\boldsymbol{\lambda}$ は MAP 推定により得る。ここで、未知パラメータ推定のための事後確率 $P(\boldsymbol{\lambda}|\mathbf{R}, \mathbf{U}, \mathbf{Q})$ は以下の式で与えられる。

$$\begin{aligned} P(\boldsymbol{\lambda}|\mathbf{R}, \mathbf{U}, \mathbf{Q}) &\propto \log(P(\boldsymbol{\lambda})) + \log(P(\mathbf{R}, \mathbf{U}, \mathbf{Q}|\boldsymbol{\lambda})) \\ &\propto -\eta \|\boldsymbol{\lambda}\|^2 + \sum_{u \in U} \sum_{q \in Q} r_{uq} \log(P(r = 1|u, q)) \\ &\quad + \sum_{u \in U} \sum_{q \in Q} (1 - r_{uq}) \log(1 - P(r = 1|u, q)). \end{aligned} \quad (8)$$

また、未知パラメータの事前確率 $P(\boldsymbol{\lambda})$ はハイパーパラメータ η により決定される平均 0 のラプラス分布に従うものとしている。式 (2) の最大化には Liu ら [5] による準ニュートン法での最適化手法等を用いる。

4. 実験

本研究では、提案手法により推定したユーザの知識、興味、嗜好で、回答する質問をどれくらいの精度で予測できるか評価する。

(注4) : <http://www.amazon.co.jp/>

4.1 データセット

本研究では、2009 年 5 月 22 日から 5 月 31 日までの間にクロールした教えて！goo のスナップショットを実験データとして用いる。上記スナップショットにおける質問数、回答数、回答者数はそれぞれ 24,272, 53,354, 5,559 であった。

4.2 評価尺度

各手法による回答する質問の予測精度を求めるのに、トップ n 正答率を用いる。トップ n 正答率は、以下のようにして求める。まず、ユーザ u の回答した質問の中でもっとも新しいものをテストデータ $\bar{q}_u$ とする。また、 $\bar{q}_u$ よりも古い質問を学習データとする。 u が回答していない質問と $\bar{q}_u$ それぞれに u が回答する確率を評価し、そのスコアがもっとも高い n 件の質問に、 $\bar{q}_u$ が含まれる場合はユーザ u を正解とみなす。このとき、全ユーザの中で正解であるユーザの割合をトップ n 正答率とする。

4.3 比較手法

ベースラインを、提案手法の素性をそれぞれ単独で用いた質問推薦とする。このとき、提案手法とベースラインは以下のようになる。

- **OurMethod.** 提案手法。ロジスティック回帰モデルにおける未知の重みパラメータ学習の際のハイパーパラメータ η は、 $\{10^{-5}, 10^{-4}, \dots, 10^4\}$ の 10 通りの候補の中から、各ユーザの回答した質問の中で最後から 2 番目の質問をテストデータとした際の (この仮のテストデータを dev データと呼ぶ) 正答率を求め、そのトップ 3 正答率を最も高くなるものとした。結果、 $\eta = 10^2$ 。
- **ContentA.** 回答履歴の本文を用いたコンテンツフィルタリング。
- **CategoryA.** もっともよく回答しているカテゴリの質問を推薦。この手法は既存のカテゴリによる質問選択とみなすことができるので、提案手法とベースラインの比較は重要である。
- **ContentQ.** 質問履歴の本文を用いたコンテンツフィルタリング。
- **CategoryQ.** もっともよく質問しているカテゴリの質問を推薦。
- **UserCF.** ユーザベースの協調フィルタリング。
- **ItemCF.** アイテムベースの協調フィルタリング。
- **Unigram.** もっとも回答されている質問を推薦する。

$$P_{\text{Unigram}}(r=1|u, q) = P(r=1|q). \quad (9)$$

4.4 結果

図 4 に、各手法のトップ n 正答率を示す。図 4 の OurMethod におけるエラーバーは各点の 99% 信頼区間を表す。まず我々は、OurMethod と、CategoryA の正答率に差があるかどうか、 χ^2 検定を行った。その際、 H_0 : OurMethod と CategoryA の正解ユーザの割合は同じであるという帰無仮説と、 H_1 : 両者の正解ユーザの割合は異なるという対立仮説を立てた。結果、有意水準 1% で全ての n について OurMethod の方が正答率が高いという結果になった。

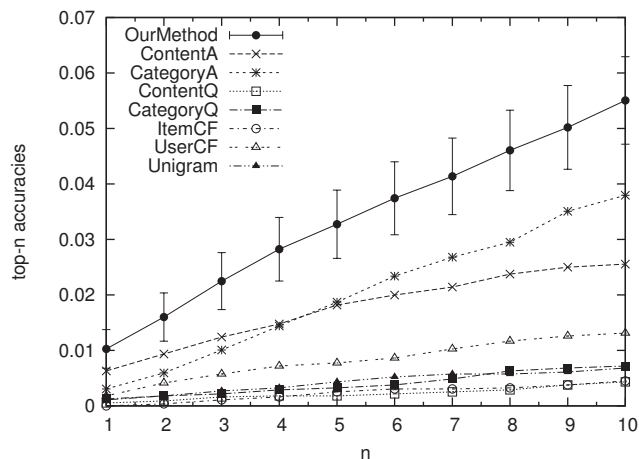


図 4 トップ n 正答率

表 3 OurMethod における各素性の重要度

素性	重要度
ContentA	5.502
CategoryA	4.396
ContentQ	0.058
CategoryQ	-0.431
UserCF	4.089
ItemCF	0.973

次に、表 3 に、OurMethod における重みパラメータ λ を示す。結果、回答履歴の本文を用いたコンテンツフィルタリングの重要度がもっとも高かった。また、カテゴリに回答する確率とユーザベースの協調フィルタリングに対する重要度が高かった。逆に、質問履歴の重要度は低かった。

4.5 考察

トップ n 正答率による推薦精度の比較実験の結果から、有意水準 1% で提案手法はカテゴリに回答する確率よりも高精度で推薦できることがわかった。すなわち、提案手法の方がカテゴリよりも正確に回答者の知識と興味に合った質問を提示できる。

OurMethod において回答履歴の方が質問履歴よりも重要度が高い。これは、重要度の学習の際、正例はユーザの回答数分与えられることが原因である。そのようなユーザに対しては、質問履歴よりも回答履歴を用いた方が精度の高い推薦が可能になる。現に実験では、質問履歴を用いた推薦よりも、回答履歴を用いた推薦の方が良い結果になった。

したがって、提案手法の重要度をそのまま使ってしまうと、回答数 0 だが質問はしているユーザに正しい推薦ができないと考える。そこで、質問履歴のみを素性とし、質問も回答もしているユーザを訓練データとして与えて学習することにより、回答数 0 の質問者に推薦できる可能性がある。

5. 分析

第 4 章にて示した予測精度の高さ以外にも、質問推薦には回答したことのないユーザにも推薦できること、回答がついていない質問であっても推薦できること、という 2 つの要件が存在する。本章では、提案法がこれら 2 つの要件を満たすかどうかについて分析を行う。

表 4 ユーザの回答数とユーザ数

1 ユーザあたりの回答数	回答数	回答者数
計	53,354	8,217
0	829	829
1	2,486	1,243
2	2,082	694
3-9	9,763	1,601
10-	38,194	1,192

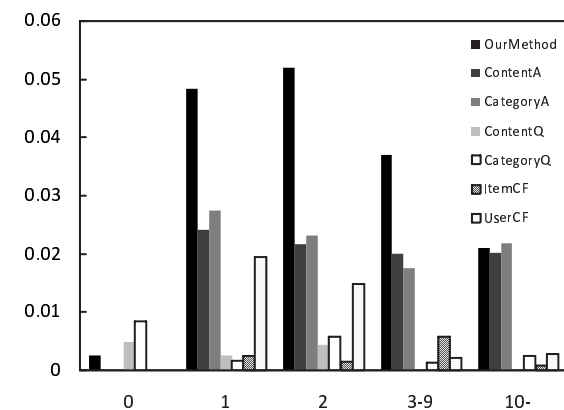
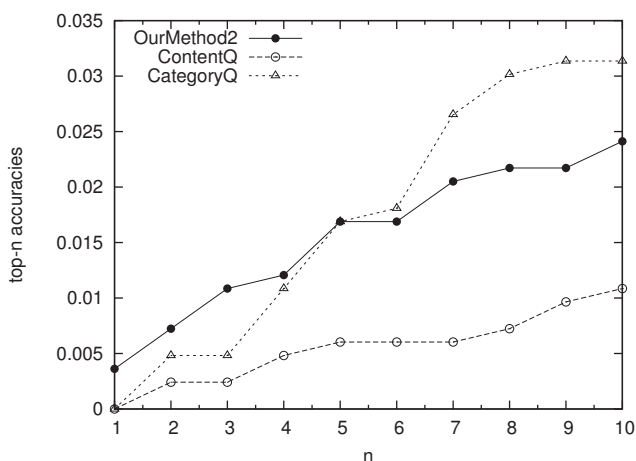


図 5 回答数で回答者をグループ分けした際のトップ 3 正答率

図 6 質問者に対するトップ n 正答率

5.1 ユーザの回答数と正答率の関係

まず、提案手法が回答したことのないユーザであっても推薦できるかどうかを検証する。表 4 に、回答数で回答者をグループ分けした際の、それぞれのグループの回答数と回答者数を示す。また、図 5 に、それぞれのグループのトップ 3 正答率を示す。提案手法に関して、回答数 1 以上の回答者であれば正しく知識と興味を推定できることがわかった。一方で、推薦する際に回答履歴を使うことのできない回答数 0 のユーザに関しては、低い正答率となった。これは質問履歴に関する素性の重みパラメータが低い値になっていることに起因する。

データセット中には回答したことはないが質問はしたことがあるユーザが 829 と約 1 割程度存在し、無視できない。しかし、質問履歴を用いれば、回答数が 0 でも質問さえしていればそのようなユーザの知識と興味を推定し、質問を推薦することは可能であると考えた。

表 5 テストデータの回答数とユーザ数

テストデータの回答数	回答者数
計	5,559
0	719
1	1,135
2	1,020
3-9	2,319
10-	366

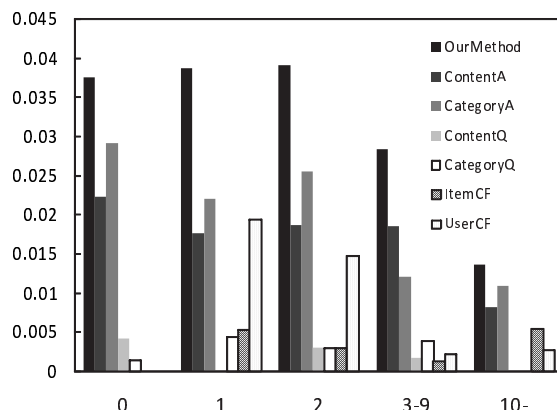


図 7 回答数で質問をグループ分けした際のトップ 3 正答率

ここで、提案法の素性を質問履歴を用いたもの（質問履歴を用いたコンテンツフィルタリングと、カテゴリに質問する確率の 2 つ）のみを用いて重みパラメータを再学習した推薦器（OurMethod2 とする）を用いる。このとき、OurMethod2 と、質問履歴による推薦手法に対して、回答数 0 の質問者を対象としたときのトップ n 正答率を図 6 に示す。 $n < 5$ においては OurMethod2 が、 $n > 5$ においては CategoryQ がもっとも有効であることがわかる。ContentQ も、比較的精度は高い。したがって、回答したことのないユーザに質問を推薦するには、そのユーザの質問履歴を用いればよいことを示すことができた。

5.2 質問の回答数と正答率の関係

次に、提案手法が未回答の質問であっても推薦できるか検証する。表 5 に、テストデータの質問についている回答数で回答者をグルーピングした際の、それぞれのグループの回答数と回答者数を示す。また、図 7 にそれぞれのグループにおけるトップ 3 正答率を示す。結果、提案手法や回答履歴を用いたコンテンツフィルタリング、カテゴリに回答する確率にもとづく手法等は質問の回答数に関わらず推薦可能であることがわかった。一方で、両協調フィルタリングは回答がついていない質問を予想通り推薦できないことがわかった。

6. 関連研究

現在、QA コミュニティに関する研究がはじょうに盛んである。Adamic ら [6] は Yahoo Answers を対象として 多角的な分析を行い、「知識共有」、「相談」、「議論」という 3 種類のコミュニケーションが行われていることを発見した。同様に、Harper ら [7] は QA コミュニティ内の知識の再利用性を高めるために、質問を “informational” なものと “conversational”

なものに分類する手法を提案した。Suryanto ら [8] は、回答者の専門性の高さを推定するアルゴリズムを提案している。Nam ら [9] は韓国の QA コミュニティサービスである Naver について、質問と回答のモチベーションに関する調査を行っている。本研究では、質問者の満足度や回答する質問を選ぶ基準など、Nam らの調査していない項目についても明らかにした。Agchitein ら [10] は、質問者が質問を投稿した際に満足するかどうかを予測するモデルを考案しており、素性に質問のタイトルや投稿時間、質問者の情報等々、QA のリッチなメタデータを活用している。Qu ら [11]、Hu ら [12] は、本研究と同様に回答者への質問推薦を提案している。我々の研究の特徴は、アンケート等の調査を通じて有効な質問推薦の要件を定義するとともに、それを満たす手法を提案した点である。また、実ログデータをつかっ、我々の提案手法がそれらの要件を満たすことを示すことができたのがこれらの研究と異なる点である。Cao ら [13] の提案した質問推薦は、本研究の質問推薦と目的が異なる。同じような内容の質問が何度も投稿されないように、質問を投稿しようとしているユーザに対して似た内容の既に投稿された質問を提示する。この質問推薦も、QA コミュニティ内の知識の再利用性を高める上で重要な機能である。

また、本研究は数多い情報推薦に関する研究の 1 つと位置づけることができる。Adomavicius ら [14] のサーベイによれば、推薦手法は大きく分けてコンテンツフィルタリングと、協調フィルタリング、両者のハイブリッド方式の 3 種類が存在する。本研究はハイブリッド方式を採用している。ハイブリッド方式を採用した既存手法としては、たとえば Jin ら [15] により提案された Web ページの推薦手法がある。また、ユーザの興味だけでなく、別の観点を考慮した推薦手法も数多く提案されている。たとえば岩田ら [16] は、購買順序を考慮し、複数のマルコフモデルを最大エントロピー原理で統合したモデルによる推薦手法を提案している。推薦の素性を吟味するのはひじょうに重要である。本研究においても、カテゴリ情報が精度の向上に大きく貢献している。

我々 [17], [18] は、QA コミュニティの質問者と回答者とのつながりをネットワークとみなし、その QA ネットワークの分析を行った。実験を通して、QA ネットワーク分析結果がたとえば「知識交換」「議論」等のコミュニケーションの種類の観点から質問を分類したり、本研究と同様回答者に対して質問を推薦したりするのに有効である可能性を示した。

7. まとめと今後の課題

本研究では、アンケート等の調査を行い、有効な質問推薦は、既存のカテゴリよりも適切な質問を提示できなければならないこと、回答したことのないユーザにも質問を推薦できなくてはならないこと、未回答の質問であっても推薦できなくてはならないこと、の 3 つの要件を満たす必要があることを明らかにした。このような有効な質問推薦を実現するために、ユーザの知識、興味、嗜好を、回答履歴、質問履歴、似た回答履歴を持つ別のユーザから推定する手法を提案した。本研究では、教えて!goo のスナップショットを用いた実験を通して提案手法が有効

であることを示した。ただし、回答したことのないユーザに質問を推薦するには、質問履歴のみを素性としてロジスティック回帰モデルの重みパラメータを再学習する必要がある。

本研究の目的は、回答の質と量の向上である。ゆえに、我々の質問推薦の手法にはまだいくつか課題が残されている。我々は以下に示す課題に取り組んでいく予定である。

- 回答の質を考慮する。質問者が良回答をつける回答者ができる回答者に、質問を推薦する手法を考える。これは、質問者が回答の質に満足でないというアンケート調査の結果からも重要である。
- 実際に QA コミュニティに我々の質問推薦を導入し、実証実験を行う。実証実験を通じて、ユーザの満足度が本当に向上するかを検証する。さらに、机上実験では検証が困難な、質問推薦の要件について検証する。たとえば、Herlocker ら [19] の提案した Novelty や Serendipity といった尺度の向上を検証する。

文 献

- [1] R.J. Mooney and L. Roy. Content-based book recommending using learning for text categorization. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pp. 195–204. ACM, 2000.
- [2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: an open architecture for collaborative filtering of netnews. In *CSCW '94: Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175–186. ACM, 1994.
- [3] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. In *WWW '01: Proceedings of the 10th international conference on World Wide Web*, pp. 285–295. ACM, 2001.
- [4] X. Jin, B. Mobasher, and Y. Zhou. A web recommendation system based on maximum entropy. In *ITCC '05: Proceedings of IEEE International Conference on Information Technology Coding and Computing*, pp. 213–218. IEEE Computer Society Department, 2005.
- [5] D.C. Liu and J. Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, Vol. 45, No. 1, pp. 503–528, 1989.
- [6] L.A. Adamic, J. Zhang, E. Bakshy, and M.S. Ackerman. Knowledge sharing and yahoo answers: everyone knows something. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pp. 665–674. ACM, 2008.
- [7] F.M. Harper, D. Moy, and J.A. Konstan. Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pp. 759–768. ACM, 2009.
- [8] M.A. Suryanto, E.P. Lim, A. Sun, and R.H.L. Chiang. Quality-aware collaborative question answering: methods and evaluation. In *WSDM '09: Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pp. 142–151. ACM, 2009.
- [9] K.K. Nam, M.S. Ackerman, and L.A. Adamic. Questions in, knowledge in?: a study of naver's question answering community. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pp. 779–788. ACM, 2009.
- [10] E. Agichtein, Y. Liu, and J. Bian. Modeling information-seeker satisfaction in community question answering. *ACM Trans. Knowl. Discov. Data*, Vol. 3, No. 2, pp. 1–27, 2009.
- [11] M. Qu, G. Qiu, X. He, C. Zhang, H. Wu, J. Bu, and

- C. Chen. Probabilistic question recommendation for question answering communities. In *WWW '09: Proceedings of the 18th international conference on World wide web*, pp. 1229–1230. ACM, 2009.
- [12] D. Hu, S. GU, S. Wang, L. Wenyin, and E. Chen. Question recommendation for user-interactive question answering systems. In *ICUIMC '08: Proceedings of the 2nd international conference on Ubiquitous information management and communication*, pp. 39–44. ACM, 2008.
- [13] Y. Cao, H. Duan, C. Lin, Y. Yu, and H. Hon. Recommending questions using the MDL-based tree cut model. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pp. 81–90. ACM, 2008.
- [14] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowl. and Data Eng.*, Vol. 17, No. 6, pp. 734–749, 2005.
- [15] X. Jin, Y. Zhou, and B. Mobasher. A maximum entropy web recommendation system: combining collaborative and content features. In *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pp. 612–617. ACM, 2005.
- [16] 岩田具治, 山田武士, 上田修功. 購買順序を効率的に用いた協調フィルタリング. 情報処理学会論文誌: 数理モデル化と応用, Vol. 49, No. SIG 4(TOM20), pp. 125–133, 2008.
- [17] 甲谷優, 川島晴美, 藤村考. QA サイトにおける質問応答グラフの成長パターン分析. 日本データベース学会論文誌, Vol. 7, No. 3, pp. 61–66, 2008.
- [18] 甲谷優, 川島晴美, 藤村考. QA コミュニティの成長パターンに基づく回答者への質問推薦. 日本データベース学会論文誌, Vol. 8, No. 1, pp. 89–94, 2009.
- [19] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, Vol. 22, No. 1, pp. 5–53, 2004.