DEIM Forum 2009 B9-5

Influential Person に着目したコミュニティ活性度評価手法の提案

佐藤　弘樹† 　島田　諭† 　福原　知宏†† 　佐藤　哲司†

† 筑波大学大学院図書館情報メディア研究科　〒 305–8550 茨城県つくば市春日 1-2
†† 東京大学人工物工学研究センター　〒 277-8568 千葉県柏市柏の葉 5-1-5
E-mail: †s0821657@ipe.tsukuba.ac.jp, ††{sat,satoh}@slis.tsukuba.ac.jp, †††fukuhara@race.u-tokyo.ac.jp

あらまし　Web 上の電子掲示板やブログ，質問回答サイトなどでは，類似した興味を持つ人々が集まりコミュニティが構築されている．これらコミュニティはそれぞれが異なる性質を持っており，その性質の形成は多種多様な要因が関係している．本研究では，コミュニティにとって影響力のある人物（IP: Influential Person）に着目することで，コミュニティを評価するための指標の提案を行う．コミュニティが活性している状態とは，ある特定の人物によって会話が成立するのではなく，より多くの人々が会話に参加している状態が望ましい．Influential Person として，本論文では Yahoo! 知恵袋のベストアンサーについて調査し，ベストアンサーを獲得する人が Influential Person として有効となるか評価した．評価した結果，ベストアンサーの存在が Yahoo! 知恵袋において質問すると，とりあえずの回答が素早く得られるという利便性に貢献していることが実データから観測できた．

キーワード　Web マイニング, Influential Person, コミュニティ活性度

An Analysis Method of Community Activities through Influential Persons on Web Communities

Hiroki SATO†, Satoshi SHIMADA†, Tomohiro FUKUHARA††, and Tetsuji SATOH†

† Graduate School of Library, Information and Media Studies, University of Tsukuba
1-2, Kasuga, Tsukuba, Ibaraki, 305-8550 Japan
†† Research into Artifacts, Center for Engineering, The University of Tokyo
Kashiwanoha 5–1–5, Kashiwa-shi, 277–8568 Japan
E-mail: †s0821657@ipe.tsukuba.ac.jp, ††{sat,satoh}@slis.tsukuba.ac.jp, †††fukuhara@race.u-tokyo.ac.jp

Abstract　Online communities, such as bulletin boards, Weblogs, and SNSs(Social Network Services) bcame popular on the WWW(World Wide Web). This paper proposes an evaluation method of Community Activities on the question and answering bulletin boards(QABBs). We measure community activities based on the behavior of the best answers who write the most useful answer in each question. Our practical evaluations using the actual QABB “ Yahoo! Chiebukuro ” figure out the index of Community Activities.

Key words　Web Maining, Influential Person, Community Activities

1. はじめに

インターネット上の電子掲示板やブログ，質問回答サイトなどでは各人の意見, それらに対するコメントを共有することで, 様々な知識の共有や新たな知識の創出がなされている．個々のサイトでは，類似した興味を持つ人々が集まり，興味に関する発言を繰り返すことでコミュニティが築かれている．インターネットに限らず実世界においてもコミュニティは栄枯を繰り返しているが，その要因は多様であり相互に依存していると考えられる．このことから，その要因を特定することは円滑なコミュニティ運営にとって重要である．ネットワークコミュニティでは投稿記事やログから様々な指標を算出でき，これらの指標を使うと運営者の主観に頼らずに的確な運営を行うことができると考えられる．

本論文では実運用で得られた Yahoo!知恵袋のデータを用いてコミュニティの活性度評価を検討する．Yahoo!知恵袋のデータを用いたコミュニティ分析には三浦 [1]，島田 [2] らの報告がある．これらの先行研究では，コミュニティの参加者全体を対象とした分析を行っているのに対して，本研究ではよりミクロに質問回答者をフォーカスすることでコミュニティの活性度を評価する．

本論文の構成は次の通りである．初めに 2. で，本論文で述べ

るコミュニティの定義と，コミュニティの活性度について基本的な考え方を示す．3. ではコミュニティを活性させる要因となるもの，Influential Person について定義を行い，次に 4. で評価手法について述べる．そして，5. で提案手法の評価を行い，6. で考察，最後に 7. で本論文のまとめと今後の課題について述べる．

2. コミュニティの活性度

Web 上の掲示板サイトやブログ，質問回答サイトでは，投稿者は各人の関心に基づいて記事を投稿し，閲覧者はその記事に興味を持った場合にコメントを行う．このことから，投稿者と閲覧者は同一の話題に対して共通の関心を持っていると考えられる．本研究では，掲示板サイトやブログ，質問回答サイトの記事やカテゴリを同じ興味を持った人たちの集まり，すなわち**コミュニティ**と捉える．

2.1 質問回答サイトにおけるコミュニティの構成

質問回答サイトでは，カテゴリ分けされた投稿スペースに質問が投稿されると，何人かの回答者が時系列上で回答をしていく．例えば Yahoo!知恵袋には現在（2009 年 2 月 20 日付）16 個のメインカテゴリがあり，質問者は自分の関心に近いカテゴリを選んで質問を投稿する．このため，個々の質問に対して図 1 に示すような複数の回答が連なった塊（本論文ではこの質問と回答集合の組を**チャンク構造**と呼ぶ）を単位として，意味的にまとまった記事集合が形成される．質問者は，自分の投稿した質問に対して満足する回答が得られた時，あるいは，これ以上の回答を望まないときに，それまでに投稿された複数の回答の中から，質問者がもっとも適切な回答であると判断した投稿をベストアンサーとして選ぶ操作を行う．このベストアンサーの決定により，その質問に対する回答が締め切られチャンク構造が確定する．

図 2 に Yahoo! 知恵袋の投稿数と投稿者数の変遷を示す．2004 年 4 月の Yahoo! 知恵袋サービス開始から 2004 年 7 月までは投稿数と投稿者数がほぼ比例して増えてきているが，2004 年 7 月から 10 月にかけて投稿者の伸びと投稿数の伸びに開きが見られるようになった．そして，2004 年 10 月以降を見ると，投稿者数は一定の割合で増えているのに対して投稿数は横ばいとなっている．投稿記事の数には伸びがなく，対して参加者は増え続けているのには Yahoo! 知恵袋のサービス開始時と以降では一人ひとりの参加者のふるまいが変わってきていると考えられる．そのため，質問回答サイトにおけるコミュニティの活性要因を見るとき，投稿数の多寡だけではなく参加者各々の役割も大きく影響していると考えられる．

2.2 質問回答サイトにおけるコミュニティの発展過程

質問回答サイトが活性化している状態とは，多くの参加者が議論に積極的に参加している状態を指す．コミュニティを活性化させるためには，そのコミュニティの現在の状態を適切に把握する必要がある．これまで，オンラインコミュニティの成功要因を分析した研究が数多く行われている [3]– [6]．これらの研究では，オンラインコミュニティの中でリーダシップを持つユーザが，オンラインコミュニティの成功と重要な関連がある

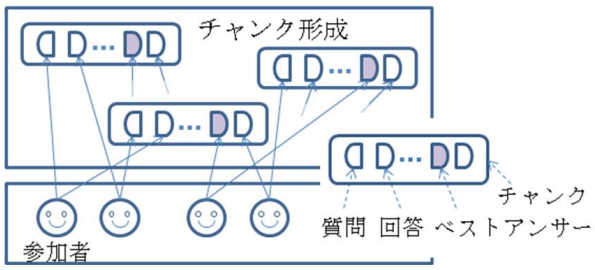

図 1　質問回答サイトのチャンク構造

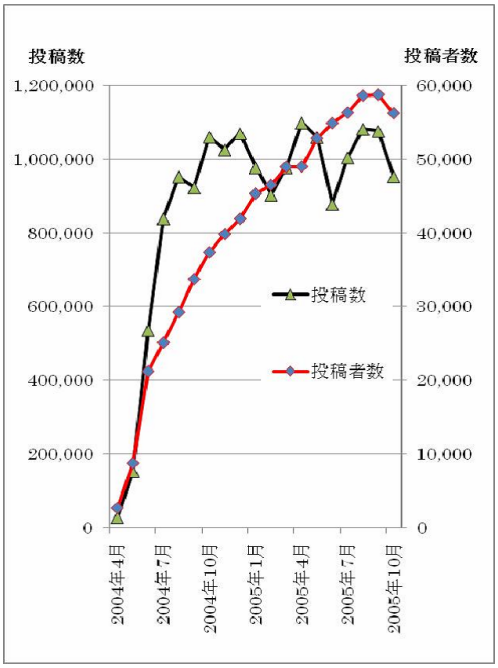

図 2　Yahoo!知恵袋の投稿数と投稿者数の変遷

ことを示している．

本論文では，コミュニティにとって影響力の大きい人に着目することでコミュニティの活性度を評価する方法を検討する．この活性度から区分されるコミュニティの性質について以下の仮説を立てた．

A)　特定の人がコミュニティを牽引していて，多くの参加者は見ているだけ（ROM（Read Only Member）状態，様子見）となっているコミュニティ

B)　多くの人が（入れ替わり立ち替わりで）発言をして，活発な議論が展開されているコミュニティ

C)　活発な議論はなされているが，発言が限定されたメンバーに限られる閉鎖的なコミュニティ

上記区分は，カテゴリの違いと時間変化の 2 つが関係すると考えられる．Yahoo!知恵袋のカテゴリによってチャンクの長さ等に違いが見られることは島田らの報告 [2] によって示されている．この区分を質問者の多寡と回答者の流動性に着目して整理すると表 1 のようになる．またこの区分は，コミュニティの発展過程における，萌芽期，成熟期，衰退期と対応させることもできる．例えば，質問回答サイトの萌芽期では参加者が増えている段階にあるため多くの人が質問者として出現するが，それに回答する人はその質問に対して見識のある特定の人に限ら

れている（表 1 の A）．しかし成熟期を迎えると，これまで質問者だった人が教えられたことにより知識を得て回答者へと移り，質問内容がこれまでと同じであったとしても回答者が特定の人からさまざまな人へと変化する（表 1 の B）．成熟期の後半では，多くの人が知識を共有していることと質問内容の飽和が起きてくることから質問数が除々に減り始めると考えられる（表 1 の B’）．衰退期では，質問回答サイトのコアメンバーが質問に対して答えるといった状態となる（表 1 の C）．

このコミュニティの発展過程は質問回答サイトに限定されない．例えば BBS や電子掲示板サイトでは，スレッドを立ち上げた人のコメントに続いて，参加者がコメントを行っている．これは質問回答サイトのチャンク構造と類似な構造となっている．そのため，BBS におけるスレッドの立ち上げ者を質問回答サイトの質問者，BBS でコメントを行う人を質問回答サイトの回答者と見なせる．したがって本論文で述べるコミュニティ発展過程は質問回答サイトに限定されるものではなく，オンラインコミュニティ全般に適用できると考えられる．

3. コミュニティの活性要因

3.1 Influential Person の定義

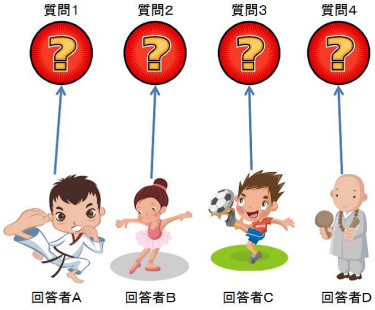

図 3　コミュニティが活性でない状態

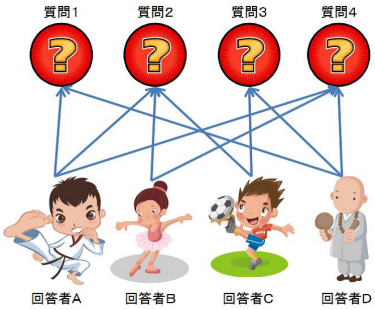

図 4　コミュニティが活性な状態

コミュニティが活性化している状態とは，多くの参加者が議論に積極的に参加している状態を指す．例えば質問回答サイトにおいて，質問に回答する人が特定の人に限られている状態（図 3）ではなく，回答する人が多様になっている状態（図 4）を指す．コミュニティの発展過程を考えたとき，投稿された質問に対してある特定の人しか答えられないというよりも，誰もがその質問に回答できるといったほうが理想的である．コミュニティを活性化した状態へと導くためには，そのコミュニティの現在の状態を適切に把握する必要がある．これまで，オンラインコミュニティの成功要因を分析した研究が数多く行われている [3]– [6]．津川らは，ユーザのソーシャルネットワークのトポロジ構造からユーザのリーダーシップ度を推定し，オープンソースコミュニティの成功度を推定する指標を提案している [6]．これらの研究では，オンラインコミュニティの中でリーダシップを持つユーザが，オンラインコミュニティの成功と重要な関連があることを示している．

コミュニティの発言を活性化させる人を発見する研究として，野村らは，掲示板のスレッドにおいて初期に投稿するユーザの行動から，将来の盛り上がりを予測する手法を提案している [7]．あるスレッドにおいて発言のバーストが見られた場合，その初期に投稿したユーザに高いスコアを与える．スコアの高いユーザの行動から，将来のバーストを予測しようとする研究である．

松村らは，掲示板における議論を主導するユーザ，すなわちオピニオン・リーダーを，他のユーザに影響を与えるという視点から Influencers として定義し，リンクと共起語を介する影響力の伝播をモデル化した手法である IDM を提案している [8]．IDM をブログに適用したところ，良好な結果が得られたと報告している [9]．

以上のことをまとめると，コミュニティの活性させる人の特徴には以下の 4 つの側面が見られる．

（1） 後続の発言を誘発するユーザ
（2） 議論を方向付けするユーザ
（3） 新しい知識をコミュニティ間で伝播させるユーザ
（4） 実世界での行動を誘発するユーザ

本論文では，以上 4 つの側面を持つコミュニティの参加者を Influential Person （IP）と定義する．松村らの Influencers は，フォーマルな議論の場を想定した（2）の側面を持つ IP だと言える．しかし，個々の投稿が自己完結しない掲示板，質問回答サイトにおいては平均文字数が短く，質問の主題語以外はほとんど共起しないことから，IDM のような手法は適用しづらい．しかし，本論文で述べる IP は環境，そして個人の専門性には依存しない．

3.2 質問回答サイトにおける Influential Person

質問回答サイトは，掲示板やブログとは異なり，一連の発言の開始と終了が，質問者によってコントロールされるという特徴がある．回答者は，サイトにアクセスした時点で回答が締め切られていない質問にのみ返信できる．このように，回答者の行動は質問の投稿状況に依存するため，個々の投稿は熟考した内容より即時的な反応になりやすい．特に，Yahoo!知恵袋では発言は 1 人 1 回に限定され，補足や追記もできない仕様になっていることから，より個々の投稿が断片化する傾向が見られる．このような環境下では，(2) 議論を方向付けするユーザ，すなわちオピニオンリーダー的な役割を果たすユーザは出現しにくいと考えられる．

また，質問者は手軽にすばやく回答が得られることを期待する傾向があり，短い場合には数十分から数時間程度で回答を締め切るユーザも少なくない．この間に回答を投稿できるユーザは自ずと限られることから，質問－回答のチャンク内で共起するユーザには，いつも同じ時間帯にサイトにアクセスしている，

表 1　コミュニティの発展過程

		質問者　の多寡	
		多数	少数
回答者の流動性	少数固定	A) **萌芽期** 参加者の多くが質問者として登場する．質問に対してある程度知識を持っている特定の参加者が回答を行う．	C) **衰退期** 質問の減少に伴い参加者も減少．新たな参加者の質問に対してコミュニティのコアメンバーが回答を行う．
	多数流動	B) **成熟期** これまで質問者だった人が特定の質問に対して回答者にもなる．そのため回答者の流動性が高まる．	B') **成熟期後半** 参加者の多くが知識をしており，誰でもが回答者になれる．一方，参加者に占める質問者が減少．

興味が似ている等，偏りがあることが考えられる．

そのため質問回答サイトにおいては議論を整理して方向付けする役割を行う IP については考慮しない．

議論を活性化し，後続の発言を誘発する役割を担う IP には，質問に対する第 1 回答者が挙げられる．先行研究の第 1 回答者を例に挙げると，第 1 回答者は後続する回答者に対して，自身が発言することで発言することへの敷居を下げる，自身の発言内容によって議論の方向性を決定づける，すなわちそのコミュニティが活性するかどうかは最初に発言する人の内容が大きい，といった影響があると考えられる．この第 1 回答者について先行研究では，全ての回答者に対する第 1 回答者の比率，第 1 回答者の回答比率，第 1 回答者の多寡の経時変化について調査を行った．その結果，Yahoo!知恵袋のカテゴリによって調査結果が異なること，時間経過によって第 1 回答者が特定の人から多数の人へと変化したことが確認できた（図 5）．

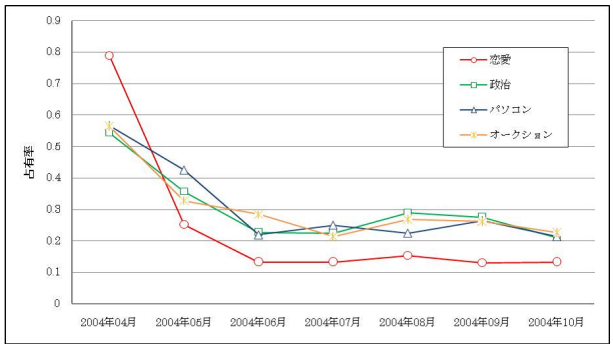

図 5　主要 4 カテゴリにおける，第 1 回答者が特定のユーザに占められる割合の月別推移

質問回答サイトにおけるコミュニティの成功とは，質問者のニーズに応えられることであるとすると，一般的な回答が短時間で得られることに加え，ありきたりでない新奇な回答（novel postings）が得られることも潜在的に求められていると考えられる．あるカテゴリにおける一般的な回答は，コミュニティの成熟により，コミュニティの内部で安定的に共有されると考えられるが，新奇な回答は外部のコミュニティから伝播するものと考えられる．

このため，3.1 節の（3）新しい知識をコミュニティ間で伝播させるユーザについて，本研究では内容的関連の必ずしも高くない多様なカテゴリにおいてベストアンサーに選ばれるユーザに着目し，これらのユーザがカテゴリ間を往来することによって，カテゴリにおける新奇な回答の出現が保証されると仮定する．コミュニティ外へのインタラクションという意味では，ベストアンサーを書いた回答者自身は他のカテゴリにおいても知識を伝搬してる可能性がある．また,（4）は本論文では扱わないが，質問者に対してもベストアンサーによって実世界に対して行動を働きかける要素も考えられる．例えば,「PC が動かなくなってしまった．どうしたらいいでしょう？」という質問に対して,「まず再起動させてみてください」という回答がベストアンサーだった場合，質問者は現実世界においてその通りに行動するなら，それはベストアンサーを書いた回答者の影響力である．本論文では，IP を複数のカテゴリにおいてベストアンサーを得るユーザとし，このユーザに対して評価を行う．

4.　コミュニティ活性度の評価手法

本論文でいうコミュニティとは，一定の分野・領域に関心を持つユーザの投稿によって形成された集合，すなわちカテゴリを指す．Yahoo!知恵袋には，細分化されたカテゴリが 286 個ある．これらの中には，内容的に重複するものや，投稿者がわずか数十人に過ぎないカテゴリも含まれる．全カテゴリを同じ尺度で扱うことは適切でなく，カテゴリをそのままコミュニティとみなすことはできない．コミュニティ活性度を評価するためには，適切な規模のコミュニティを抽出する必要がある．本論文では，カテゴリ間で共起するユーザ数に基づき，カテゴリ間の関連度を求め，関連度の高いカテゴリ群を一つのコミュニティとみなす．また，質問回答サイト特有の構造として，ユーザの投稿が質問と回答に厳密に分かれており，またベストアンサーを選ぶ仕組みも有しているため，これらの投稿区分ごとにカテゴリ間の関連度を求める．例えば，カテゴリ A でベストアンサーを出す人は IP の候補者 IP(A) であるとする．次に，その人がカテゴリ B でもベストアンサーを出せば IP(A,B) である．IP(A,B) はカテゴリ A,B 間をブリッジする役割を果たす．また IP(X,Y) が高いカテゴリ X と Y は近似したカテゴリであるということができると考えられる．

ベストアンサーの選択基準は，質問者が納得した回答につけ

表 2　分析対象カテゴリの概要

カテゴリ名	USER	USER_q	USER_ba	USER_na	POST_q	POST_na
恋愛相談、人間関係の悩み	68,199	24,093	30,867	59,228	210,105	1,206,457
Yahoo	50,831	17,734	25,409	42,297	208,793	926,532
Yahoo!オークション	51,054	28,434	16,258	35,451	190,432	1,052,214
パソコン、周辺機器	58,102	43,494	14,550	23,420	171,848	302,839
病気、症状、ヘルスケア	53,999	29,954	19,421	32,926	84,394	205,184
政治、社会問題	30,978	13,259	11,896	23,570	78,777	324,529
テレビ、ラジオ	40,810	20,895	17,551	27,513	75,194	169,270
インターネット	40,078	25,965	12,948	17,981	70,529	102,949
芸能人、タレント	36,404	14,639	14,476	27,313	55,783	189,946
レシピ、調理法	35,850	17,232	11,225	25,095	52,430	197,498
健康、病気、ダイエット	35,265	15,259	13,868	22,489	41,978	95,441
料理、グルメ、レシピ	29,264	11,959	10,125	20,681	37,714	123,526
子育ての悩み	27,170	8,164	8,741	22,620	35,153	182,973
携帯電話、モバイル	26,534	15,419	8,187	12,980	35,001	56,884
話題の人物	30,726	10,245	11,227	24,158	34,916	146,758
ゲーム	20,241	11,233	6,970	10,772	30,085	51,391
マナー	25,793	9,352	7,625	19,626	23,693	127,581
家電、AV 機器	20,220	11,677	5,267	9,137	23,488	45,611
あの人は今	17,293	6,201	5,690	12,150	15,589	44,481
法律相談	16,610	7,667	4,140	9,482	15,513	38,753
飲食店、デパ地下	16,333	6,514	5,386	10,109	12,158	30,297
子どもの病気とトラブル	11,411	3,645	3,123	8,333	7,491	27,701
これ、探してます	10,348	4,309	3,546	4,793	6,087	8,358
あいさつ、てがみ、文例	9,686	3,274	2,267	6,259	4,359	16,482

ることになっている．これは回答が，質問に対する正解であれば必ずしもベストアンサーに選ばれるわけではない．Yahoo!知恵袋においては一概に正解が決められない質問が多く，質問者は正解や不正解という観点よりも，回答に対する納得度，もしくは感謝という点でベストアンサーを選ぶ傾向が見受けられる．そのため，質問者を納得させられる回答を書くことがベストアンサーに最も近づく方法となるが，それには習熟が必要と考えられる．そこで，ベストアンサーを連発する人というのはカテゴリに特化した専門性よりも，Yahoo! 知恵袋全体に通用するような回答作成能力、すなわち質問者が納得しやすい回答を書ける能力を持つと仮定する．

5. 評　　価

5.1 調 査 方 法

4. 節の仮定を検証するため，以下の 2 種類の調査を行った．

（1） ベストアンサーとノーマルアンサーの文章の計量的指標（文字数）の比較

（2） ベストアンサーに選ばれる人の出現カテゴリ数

(1) については，エビデンスとなる内容が多く盛り込まれた客観的な回答がベストアンサーに選ばれやすいとすると，ベストアンサーはノーマルアンサーに比べて文字数が多いという相関が見られると考えられる．(2) については，ベストアンサーに多く選ばれている人が，内容の関連するカテゴリ群に留まらず，多種多様なカテゴリでベストアンサーに選ばれていれば，ベストアンサーの作成能力と，カテゴリに特化した専門性との相関は低いとみなすことができる．

調査対象は，Yahoo!知恵袋データの全期間（2004 年 4 月〜2005 年 10 月）を対象に分析を行う．特記のない限り，全カテゴリ・全記事を対象とする．その中から（1）の調査では，カテゴリ別投稿数上位 4 つ「恋愛相談、人間関係の悩み」「政治、社会問題」「パソコン、周辺機器」「Yahoo! オークション」（以下,「恋愛」「政治」「パソコン」「オークション」と略）から選択,（2）ではその 4 カテゴリに関連する 20 カテゴリを加えた計 20 カテゴリを調査した（表 2）．文字数は MySQL の CHAR_LENGTH() 関数を用いてカウントした．マルチバイト文字は 1 文字として数える．

5.2 評 価 結 果

5.1 節の (1) について，全カテゴリ全期間の文字数の分布をベストアンサー，ノーマルアンサー別に調べた．その結果，図 6–図 8 となり，いずれも 30 文字を最頻値とする対数正規分布

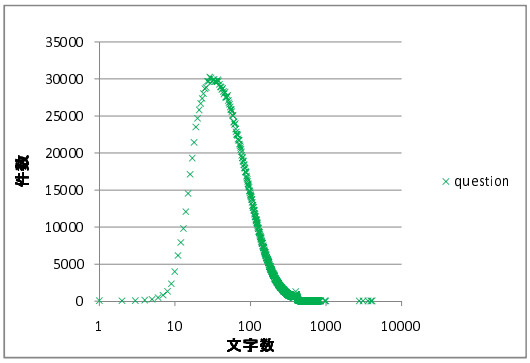

図 6　全期間・全カテゴリの質問における文字数の分布

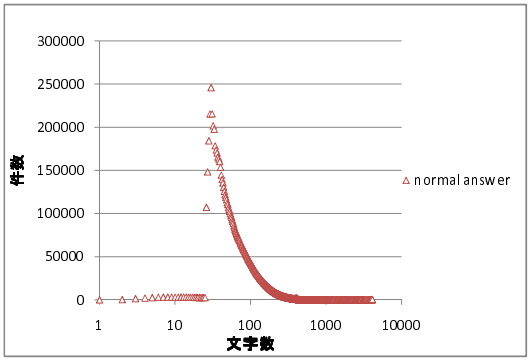

図 7　全期間・全カテゴリのノーマルアンサーにおける文字数の分布

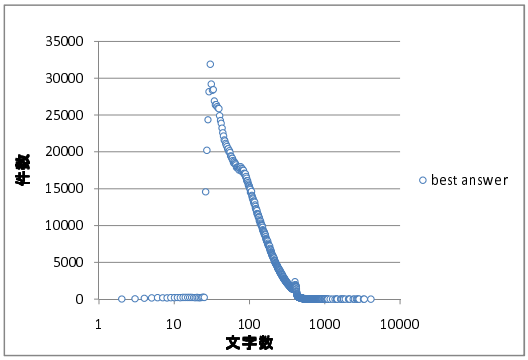

図 8　全期間・全カテゴリのベストアンサーにおける文字数の分布

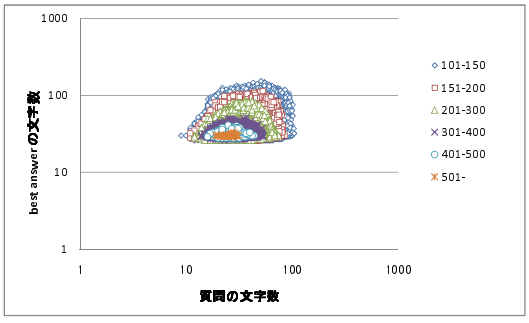

図 9　全カテゴリにおける質問とベストアンサーの文字数の相関図

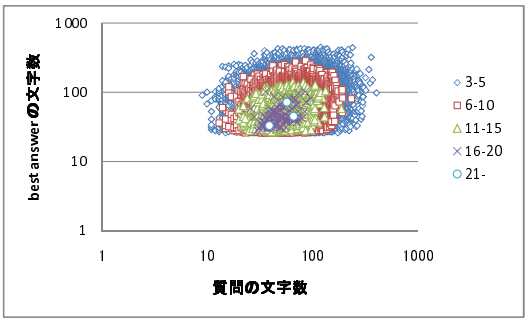

図 10　PC カテゴリにおける質問とベストアンサーの文字数の相関図

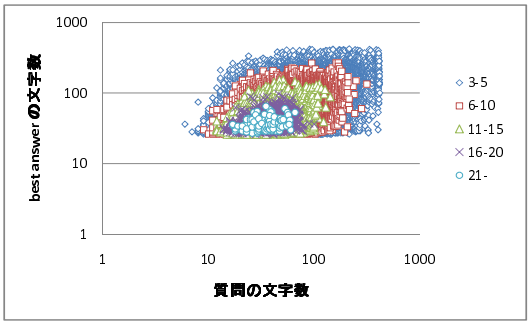

図 11　恋愛カテゴリにおける質問とベストアンサーの文字数の相関図

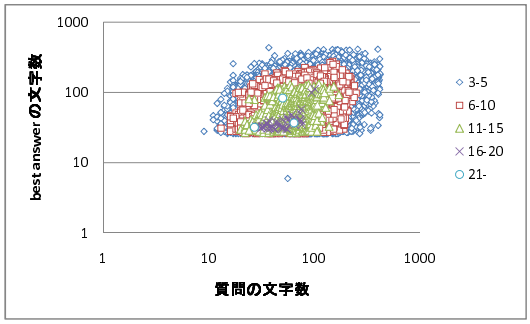

図 12　オークションカテゴリにおける質問とベストアンサーの文字数の相関図

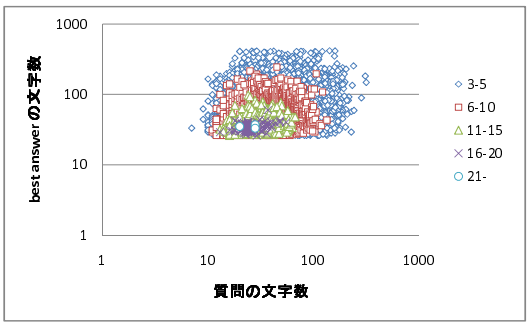

図 13　政治カテゴリにおける質問とベストアンサーの文字数の相関図

となっていた．ただし，ベストアンサーとノーマルアンサーは最低文字数が 25 文字以上という知恵袋の制約のため，分布の前半が欠けた形になっていた．また，全カテゴリと今回注目する 4 大カテゴリについて質問の文字数とベストアンサーの文字数の相関も調査した．質問とベストアンサーの相関を，図 9–図 13 に示す．結果から質問とベストアンサーの文字数の相関係数は 0.18 となり，ほとんど相関はない．

5.1 節の（2）について，質問数の多い主要な 4 カテゴリと，質問数が一定以上であり主要カテゴリに関連すると思われる 20 カテゴリを加えた計 24 カテゴリについて，各 2 カテゴリ間において両方のカテゴリでそれぞれ 1 回以上ベストアンサーに選ばれているユーザの人数（ブリッジ人数）を調べた．その結果を，図 14 に示す．調査結果から，ほぼ全てのカテゴリ間でベストアンサーに選ばれたことのあるユーザが 2000 人以上となっていた．このことから，カテゴリ間の相関は高く，あらゆるカテゴリに出現しベストアンサーを獲得するユーザの存在が観測できた．また,「恋愛相談、人間関係の悩み」と「Yahoo! 知恵袋」など，ブリッジ人数が 10000 人以上となるカテゴリの組み合わせが存在した．それに対して,「携帯電話、モバイル」と「法律相談」のようにブリッジ人数が少ないカテゴリの組み合わせも見られた．このように，ベストアンサーのブリッジ人数が多くなりやすい，または少なくなりやすいといったカテゴリの組み合わせによる傾向の差が観測できた．

6. 考　　察

調査結果から，ベストアンサーを出す人はノーマルアンサーと比較し文字数に差がないことから，書く文章の性質はあまり変わらないことがわかる．次に質問とベストアンサーの文字数を調査した結果，質問が長いから回答も長くなるというような性質も特別に見られるわけではなかった．また，ベストアンサーを極端に連発する人は，ベストアンサーの獲得を狙って書いている稀有な人であると懸念されている．実際,「ベストアンサー」に言及しているユニークユーザ数はわずか数% に留まることから，そのような特殊な人の割合は十分に小さいと考えられる．そのため，ベストアンサーとなる回答者は質問に対して自由に回答を行った結果，偶然ベストアンサーとなっていると推測される．

しかし，調査（2）の結果から，カテゴリをまたいでベストアンサーを獲得する人が多々存在することがわかった．このことからベストアンサーとなる回答者は偶然ベストアンサーに選ばれてはいるが，一つのカテゴリに依存せずあらゆるカテゴリにおいて，質問者を納得させる回答を行うことができる人であると考えられる．このようにベストアンサーの存在が Yahoo! 知恵袋で質問すると，回答が確実に得られるという利便性に貢献していることが実データから観測できた．また，今回あるカテゴリ間においてブリッジする人数が多く見られる傾向が確認できた．これには,（1）factoid 型や non-factoid 型といったそのカテゴリの性質,（2）カテゴリ内でされている質問内容の近さ，などが考えられるが，ここをどう評価するかは今後の課題である．

7. まとめと今後の課題

インターネット上の掲示板サイトやブログ，質問回答サイトなどで同じ興味をもった人たちの集まりをコミュニティと定義し，そのコミュニティの活性度評価を行うための検討を行った．本研究では，コミュニティの活性度に影響を与えるユーザを Influential Person として，この Influential Person に着目することでコミュニティ活性度の評価する手法を提案した．インターネット上で実運用された質問回答サイト (Yahoo! 知恵袋) のアーカイブデータを用いて調査を行ったところ，ベストアンサーが Influential Person として有効であることがわかった．

今後の課題として，あらゆるカテゴリにおいてベストアンサーを獲得する人の出現するカテゴリ間の関係及び性質の調査，ベストアンサーとコミュニティの発展過程との相関を分析を行い，コミュニティの活性度を評価する指標としてモデル化を行う．

謝　　辞

本研究の実施にあたっては，ヤフー株式会社が国立情報学研究所に提供した Yahoo! 知恵袋データを利用した．ここに記して謝意を示す．

文　　献

[1] 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，第 20 回人工知能学会全国大会，3D3-1，2006.

[2] 島田諭，福原知宏，佐藤哲司：質問回答サイトにおける利用者行動に基づく記事の関連付け手法の検討，第 19 回データ工学ワークショップ (DEWS2008)，B6-5，2008.

[3] 松村憲一，西田豊明: コミュニケーションツール評価手法の構築, 社会技術研究論文集 Vol.2，pp.181-190, 社会技術研究会，2004.

[4] A. Bourhis，L. Dube，and R. Jacob：The success of virtual communities of practice: The leadership factor，Electronic Journal of Knowledge Management，vol. 3，pp.23–34，2005.

[5] H. Lin，W. Fan，L. Wallace，and Z. Zhang：An empirical study of web-based knowledge community success，Proceedings of the 40th Annual Hawaii International Conference on Information Systems(HICSS ’07)，p.178，2007.

[6] 津川 翔，杉山 浩平，大崎 博之，今瀬 真：ソーシャルネットワークのトポロジ構造を用いた開発型オンラインコミュニティの成功度推定手法，電子情報通信学会技術研究報告 (IN2008-61)，pp. 105-110，2008.

[7] 野村拓也，小野稔，下山洋一，高木友博：目利き人を利用した電子掲示板の盛り上がり予測，Web インテリジェンスとインタラクション研究会（WI2），2008.

[8] Naohiro Matsumura, Yukio Ohsawa, and Mitsuru Ishizuka: Influence Diffusion Model in Text-based Communication, WWW02, 2002.

[9] Naohiro Matsumura, Hikaru Yamamoto, Daisuke Tomozawa: Finding Influencers and Consumer Insights in the Blogosphere, International Conference on Weblogs and Social Media (ICWSM), Seattle, WA, March 31-April 2, pp. 76–83, 2008.

[10] 佐藤弘樹, 島田諭, 福原知宏, 佐藤哲司: コミュニティの活性度評価に関する一検討, 情報社会学会知識共有コミュニティワークショップ論文集, pp.29-38，2008.

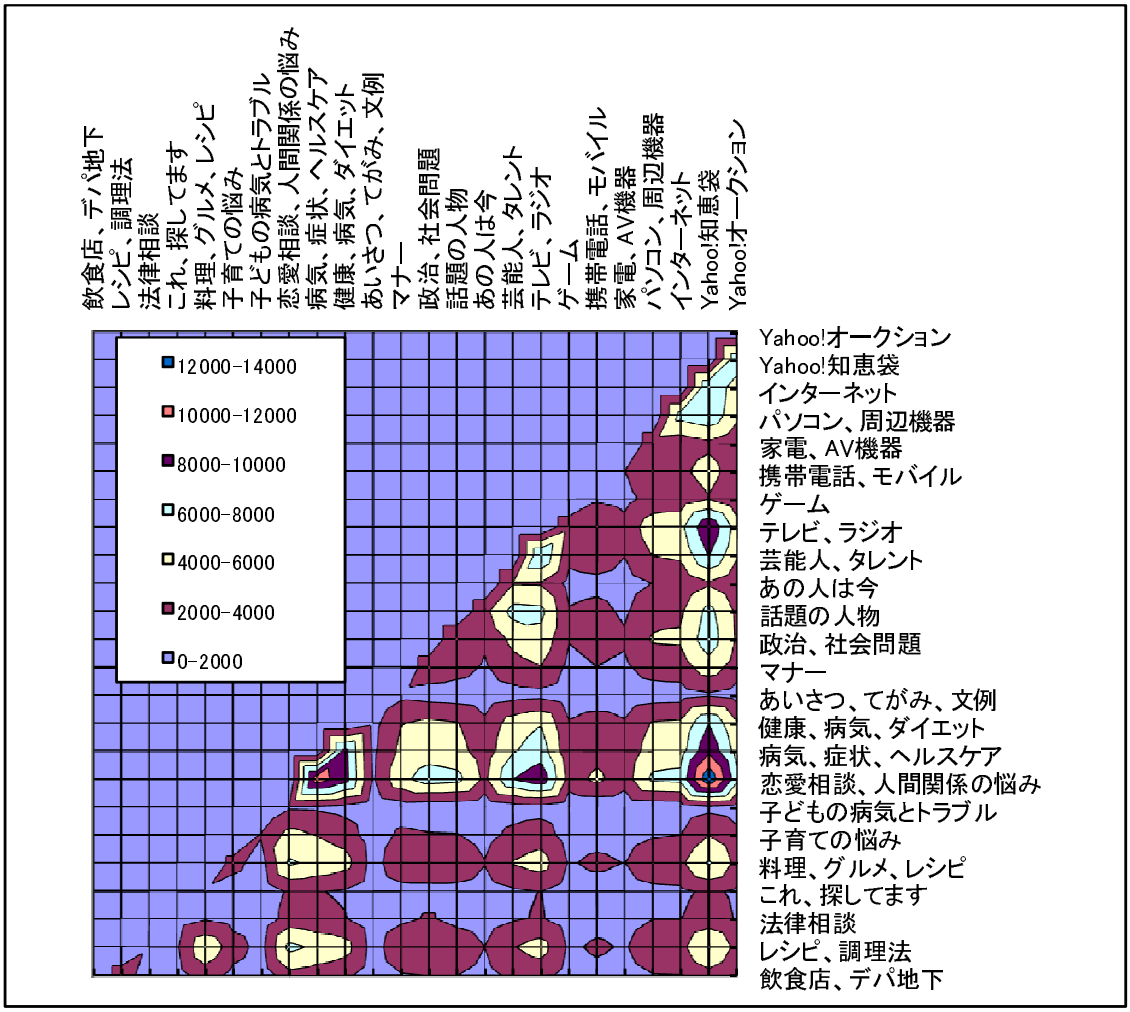

図 14　各 2 カテゴリ間において両方のカテゴリでベストアンサーに選ばれているユーザの人数（ブリッジ人数）