

情報共有を基盤とするソーシャルメディアサイトの
活性化に関する研究

—Q&A サイトの報酬制度設計と Novelty 向上のための推薦手法の提案—

小川 祐樹

電気通信大学

2011年3月

情報共有を基盤とするソーシャルメディアサイトの
活性化に関する研究

—Q&A サイトの報酬制度設計と Novelty 向上のための推薦手法の提案—

小川 祐樹

電気通信大学大学院情報システム学研究科

社会知能情報学専攻

博士（工学）学位申請論文

情報共有を基盤とするソーシャルメディアサイトの 活性化に関する研究

—Q&A サイトの報酬制度設計と Novelty 向上のための推薦手法の提案—

博士論文審査委員会

主査	太田敏澄	教授
委員	藤村考	教授
委員	大須賀昭彦	教授
委員	田野俊一	教授
委員	本多弘樹	教授

著作権所有者

小川 祐樹

2011年3月

Study for Activity of Social Media Site based on Information Sharing: The Design of the Payback System of a Q&A Site, and the Proposal of a Recommendation System with Novelty

Yuki Ogawa

Abstract

In this study, an effective design policy is proposed from a viewpoint of 2 systems of information accumulation and information processing to the problem of a design of the social media site based on information sharing. About the system of information accumulation, the payback system which is one of the motivation support of an information provider showed that the information accumulated at a Q&A site was affected, and clarified the effective design policy in the number of answers, and a nature viewpoint. About the system of information processing, the recommendation technique for the improvement in Novelty which supports the contact to various information was proposed, and the subject experiment showed the usefulness. The design policy proposed by this study is new knowledge which supports the function, and the institutional introduction and the employment in a social media site based on information sharing.

情報共有を基盤とするソーシャルメディアサイトの活性化に関する研究 —Q&A サイトの報酬制度設計と Novelty 向上のための推薦手法の提案—

小川 祐樹

論文要旨

本論文では、情報共有を基盤とするソーシャルメディアサイトの機能・制度設計の問題に対して、情報蓄積と情報加工の 2 つの設計課題を挙げ、有効な設計方を提案することを目的とする。個々の参加者間の情報受発信行動によって新たな価値が生成されるソーシャルメディアにおいて、参加者の情報発信を促す情報蓄積のための設計や、その蓄積された情報を有用なカタチで参加者に提供するという情報加工の設計は、参加者に有用な価値を提供し続ける持続発展可能なソーシャルメディアの実現において重要な課題である。

第一の課題である情報蓄積の設計課題に関しては、参加者主体の知識共有が行われる Q&A サイトをとりあげる。ここでは、知識提供者である回答者のモチベーション支援制度のひとつである報酬制度について、効果的な情報蓄積のための報酬制度設計の課題に取り組む。従来、知識提供を行う回答者のモチベーションに関しては、援助的動機、社会的動機、報酬的動機、互酬的動機が存在することが指摘されている。しかし、これらの動機をどのように支援するかといった制度設計に関する知見や、どのような動機支援が参加者間の相互作用に影響をあたえ、結果的にどのような知識生成をもたらすかといった効果に関する客観的な知見は得られていない。ソーシャルメディアのひとつである Q&A サイトにおいて、このようなモチベーション支援の方策や、その効果を理解して効果的な運用方針を設定することは、情報を必要とする質問者を助けると同時に、サイト上での新たな情報蓄積を促し、他の参加者にとって有益な情報を提供することに繋がる重要な課題である。

以上をふまえ、本論文における情報蓄積の課題では、知識提供者のモチベーション支援のひとつである報酬動機に着目し、報酬制度設計が質問・回答の数や質に与える影響を分析することを目的とした研究課題に取り組んだ。具体的には、Q&A サイトにおける質問者・回答者の知識取引をマルチエージェントシステムとしてモデル化し、シミュレーション実験によって報酬制度がもたらす効果の分析を行った。その結果、提案したモデルに関して、提案モデルでのエージェントの利用行動と現実の回答者の利用行動との類似性を確認し、報酬制度の効果に関しては、そのコミュニティが何を重視するのか（回答の数、回答の質、質問のしやすさ）によって有効な報酬制度は異なってくるということを明らかにした。また、報酬制度の運用方針として、回答の数を重視するのであれば質問者と回答者との知識取引における報酬は低く、回答の質と質問のしやすさを重視するのであれば報酬は高く設定するという運用方針に関する知見を明らかにした。さらに、質問者と回答者の知識傾向の違いといったコミュニティ環境ごとのシミュレーションによる分析によって、コミュニ

ティ環境ごとの効果的な報酬制度の運用方針の知見を明らかにした。

つぎに、第二の課題である情報加工の設計課題に関しては、個々の参加者の多数の行動履歴を加工することによって実現されるソーシャルフィルタリング機能の設計課題について取り組んだ。膨大な情報の中から、利用者の望む情報を加工して提供してくれる推薦システムは、現在多くのニュースサイトや EC サイトなどで用いられてきている。特に、他者の意見を利用するソーシャルレコメンドの機能は、多数の個人の断片的な行動データから、嗜好傾向や社会的トレンドといった個人の個々の情報だけでは得られなかった新たな価値ある情報の生成を可能にしている。このような機能は、ソーシャルメディアにおける多数の参加者からの情報生成を支援する重要な機能のひとつである。しかし、従来の推薦に関する研究において、情報推薦の有用性については、予測精度の適切さといった点に基づいて評価されることが多く、Novelty（未知かつ好みのコンテンツの発見）の観点での推薦手法に関する研究は少ない。いくつかの手法が検討されているものの、利用者の労力の高さや、コンテンツの特性に限定されてしまうなど、適用範囲の広い手法に関しては十分に検討されていない。推薦における Novelty は、ユーザの既存の嗜好・興味範囲の過度な固定化を防ぎ、推薦への飽きの解消や、興味の開拓を促す重要な要素である。これは、有用な価値を提供できる情報推薦（ユーザの長期的な利益を考慮した推薦）を実現するうえでの重要な設計課題である。

以上をふまえ、情報加工の課題では、利用者の労力が少なく、かつコンテンツの適用範囲が広い Novelty 向上の推薦手法を提案することを目的とした研究課題に取り組んだ。ここでは、嗜好傾向が類似するアイテムを複数推薦するよりも、嗜好傾向の異なるコンテンツを複数推薦する方が、利用者が未知のコンテンツを推薦する可能性が高まるという仮定を立て、トピック多様化によって Novelty 向上を行うアプローチを提案した。具体的には、利用者のコンテンツに対する評価情報といった行動履歴のデータのみを用いて、嗜好傾向が類似するコンテンツをクラスタリングによりトピックとして自動抽出した。さらに、このトピック情報と既存の推薦リストの推薦順位とを用いて、多様なトピックのコンテンツを選定することで、嗜好傾向が多様な推薦リストを作成する手法を提案した。そして、評価実験として、DVD コンテンツを用いて多様化を行った推薦リストをユーザに提示し、アンケート調査により提案手法の Novelty 評価を行った。実験の結果より、提案手法が、既存の推薦手法である協調フィルタリング手法と比較し、Novelty を向上させる手法であることを確認した。これにより、利用者のコンテンツに対する評価情報というデータのみを用いた適用範囲の広い Novelty 向上の推薦手法を実現した。

結論として、本論文では、情報共有を基盤とするソーシャルメディアサイトの情報蓄積と情報加工の設計課題をとりあげ、報酬制度と Novelty 推薦に対して有効な機能・制度設計を示した。本論文で提案した設計は、実際の Web サービスにも応用可能であり、参加者に対し新たな情報を提供し続ける持続発展可能なソーシャルメディアの構築を支援するためのひとつの有効なアプローチであるといえる。

目次

1 章. 序論.....	1
1.1. 研究背景.....	1
1.1.1. ソーシャルメディア	1
1.1.2. ユーザの有用な情報取得を支援する情報共有サイト	2
1.1.3. 情報共有サイトの活性化.....	3
1.1.4. 情報共有サイトを支える「情報蓄積」と「情報加工」のシステム	3
1.1.5. 情報共有サイトにおけるユーザの有用な情報取得の課題.....	5
1.2. 本研究の目的と意義.....	8
1.2.1. 本研究の目的	8
1.2.2. 本研究の意義.....	8
1.3. 本論文の構成	9
2 章. 情報共有サイトの支援技術.....	11
2.1. 情報共有サイトにおける情報提供のモチベーション支援技術	11
2.1.1. 情報提供のモチベーション	11
2.1.2. 情報提供の外発的モチベーション	12
2.1.3. 情報共有サイトにおける報酬制度の重要性.....	13
2.1.4. Q&A サイトにおける報酬制度	14
2.2. 情報共有サイトにおける有用な情報の取得支援技術.....	14
2.2.1. 情報提供の形態	14
2.2.2. 情報共有サイトにおける情報推薦システムの重要性.....	15
2.2.3. 他者意見を利用した情報推薦システム	16
2.2.4. 情報共有サイトにおける情報推薦システムの課題	17
2.3. 本研究の位置づけと課題.....	18
2.3.1. 情報提供のモチベーション支援	18
2.3.2. 有用な情報の取得支援	19
3 章. Q&A サイトにおける効果的な報酬制度設計.....	21
3.1. 研究背景.....	21
3.2. 先行研究と本研究の位置づけ	22
3.2.1. Q&A サイトに関する研究.....	22
3.2.2. エージェントベースシミュレーションによるに関する研究.....	23
3.2.3. 本研究の位置づけ.....	24

3.3.	Q&A サイトのエージェントベースモデル	25
3.3.1.	知識取引のモデル化	25
3.3.2.	利用者のモデル化	26
3.3.3.	報酬制度のモデル化	30
3.3.4.	利用者の効用のモデル化	31
3.3.5.	コミュニティ環境（質問者と回答者の同質性）	32
3.3.6.	利用者の離脱・参入	32
3.4.	シミュレーション実験	33
3.4.1.	実験概要	33
3.4.2.	モデルの基本的挙動	35
3.4.3.	報酬制度の導入効果	37
3.5.	考察	42
3.5.1.	回答の数と質	42
3.5.2.	質問努力（質問のしやすさ）	43
3.5.3.	報酬制度の導入方針への知見	44
3.5.4.	本シミュレーションの限界と制約	45
3.6.	結論と今後の課題	45
4 章.	Novelty 向上のための推薦手法の設計	47
4.1.	研究背景	47
4.2.	先行研究と本研究の位置づけ	48
4.2.1.	従来の推薦システムのアルゴリズムと評価	48
4.2.2.	推薦の正確さ以外の向上を目的とした研究	50
4.3.	動的なトピック分類に基づく推薦リストのトピック多様化推薦	52
4.3.1.	アルゴリズムの概要	53
4.3.2.	嗜好傾向に基づくアイテムのトピック分類（Step 1）	54
4.3.3.	ユーザごとにパーソナライズされた推薦リストの作成（Step 2）	55
4.3.4.	トピックを用いた推薦リストの多様化（Step 3）	56
4.4.	評価実験	58
4.4.1.	実験概要	58
4.4.2.	実験結果	60
4.4.3.	多様化推薦リストの結果	64
4.4.4.	ユーザ満足度評価の結果	66

4.5.	考察.....	68
4.6.	結論と今後の課題.....	69
5章.	考察.....	71
5.1.	研究成果のまとめ.....	71
5.2.	研究の展望と限界.....	72
6章.	結論と今後の課題.....	75
6.1.	結論.....	75
6.2	今後の課題.....	76
	謝辞.....	78
	参考文献.....	79
	関連論文の印刷公表の方法及び時期.....	86

図目次

図 1-1 : 従来のメディアとソーシャルメディアでの情報の生成・提供の違い	2
図 1-2 : 情報共有サイトにおける「情報蓄積」と「情報加工」のシステム	4
図 1-3 : 情報共有サイトの活性化.....	8
図 1-4 : 本研究の構造.....	10
図 3-1 : モデル上の回答者と、現実の Q&A サイトの回答者の特性の比較	35
図 3-2 : 報酬 PT を変化させたときの、回答の数の変化	37
図 3-3 : 報酬 PT を変化させたときの、回答の質の変化	38
図 3-4 : 報酬 PT を変化させたときの、質問努力の変化	39
図 3-5 : 同質性 $z=0.5$ (質問者と回答者の同質性が中間のコミュニティ環境) における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係..	40
図 3-6 : 同質性 $z=0.3$ (質問者と回答者の同質性が中間のコミュニティ環境) における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係..	41
図 3-7 : 同質性 $z=0.7$ (質問者と回答者の同質性が中間のコミュニティ環境) における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係..	41
図 3-8 : 同質性 $z=0.3$ と同質性 $z=0.7$ のコミュニティ環境での、報酬 PT=0.0 における質問努力と回答数の関係の違い.....	42
図 4-1 : 提案する推薦システムの全体図	54
図 4-2 : トピック多様化のアルゴリズム	57
図 4-3 : アイテムの共評価ネットワーク (ノード数 : 1,000 エッジ数 : 9,171)	61
図 4-4 : クラスタリング結果 (ノードの大きさはクラスタ内ノード数の多さを表す)	61
図 4-5 : 推薦リストの上位 X アイテム中における トピック網羅数	66
図 4-6 : Novelty : 推薦リストの上位 X アイテム中における「知らなかったが、推薦されて興味を持ったアイテム」の割合	67
図 4-7 : Accuracy : 推薦リストの上位 X アイテム中における「知っている・好みのアイテム」の割合	67
図 4-8 Precision : 推薦リストの上位 X アイテム中における「好みのアイテム」の割合	68

表目次

表 1: シミュレーション実験パラメータ (表中の Q は質問者、A は回答者を表す) }	34
表 2: 2007 年次のトピック分類結果 (一部)	62
表 3: 2007 年次と 2008 年次のトピックの変化 (一部)	63
表 4: CF でのリスト (左) と提案手法でのリスト (右) の比較	65

1章. 序論

1.1. 研究背景

1.1.1. ソーシャルメディア

近年、多数のユーザの情報受発信行動によって価値ある情報を発展的に生成していく「ソーシャルメディア」と呼ばれる新たなメディアが浸透しつつある。ソーシャルメディアとは、広くはユーザの情報受発信によって UGC (User Generated Content) の生成を可能とする Web サイトの総称とされており [Kaplan 10]、ブログや SNS、Twitter、Q&A サイト、レビューサイト、Wiki など多様な Web サイトとして実現されている [Yoshikawa 09, Kaplan 10]。

ソーシャルメディアの特徴は、ユーザの情報受発信によって、価値ある情報が発展的に生成される点である。ソーシャルメディアを実現する Web サイトにおいて、ユーザは自身の意見や経験談の発信が可能であり、これらのユーザが発信する情報は Web サイトという場に蓄積される。蓄積された情報は、日常の出来事に対する感想をつづった日記的なものや、購入した商品に対するレビュー、または自身の持つ経験や知識を書きつづったものなど多様な情報であり、これを閲覧する他者にとって有用な情報源となっている。例えば、ユーザがブログやレビューサイトに書き込む商品レビューや、商品の使用感などの経験に基づく意見や評価の情報は、他のユーザにとって商品購入の意思決定を行う際の重要な情報源になっている [山本 09]。さらに、ソーシャルメディアでは、ユーザは他者が発信した情報に対してコメントや、評価・投票が可能であり (Wikipedia、Q&A サイト、レビューサイト等)、このような情報受発信行動は、提供された情報の内容の充実や洗練や集合知 [Surowiecki 06] をもたらし、ユーザによる発展的な情報生成を可能としている。

このような特徴を持つソーシャルメディアは、従来のマスメディア (TV、新聞、ラジオ) が行う情報の生成や提供のかたちとは大きく異なっている点が特徴的である。従来のマスメディアにおいては、情報の生成を行う特定少数の送り手が、消費者である個人に対して一方向的な情報・コンテンツ提供を行っていた。しかし、ソーシャルメディアにおいては、個人はブログやレビューサイトなどの Web サイトを通じた自身の意見や経験談の発信、または提供されている情報へのレビューによって、ユーザ主体の情報生成を可能にしている (図 1-1)。また、ソーシャルメディアの利用は、個人の情報取

得といった利用にとどまらず、多くの分野での活用が行われている。例えば、地震などの災害時におけるマスメディアを補完する新たな情報源としての活用 [垂水 10, Sakaki 10]、企業内での情報共有 [Li 08, Nexti 運営メンバー有志 10] や、新たな CRM ツールとしての可能性 [Li 08]、あるいは、政治や地域行政への利用用途 [NPO 法人ドットジェイピー 06, 庄司 10, 地方自治情報センター 07, 庄司 08] としても注目され、ソーシャルメディアの社会的な影響は大きくなりつつある。

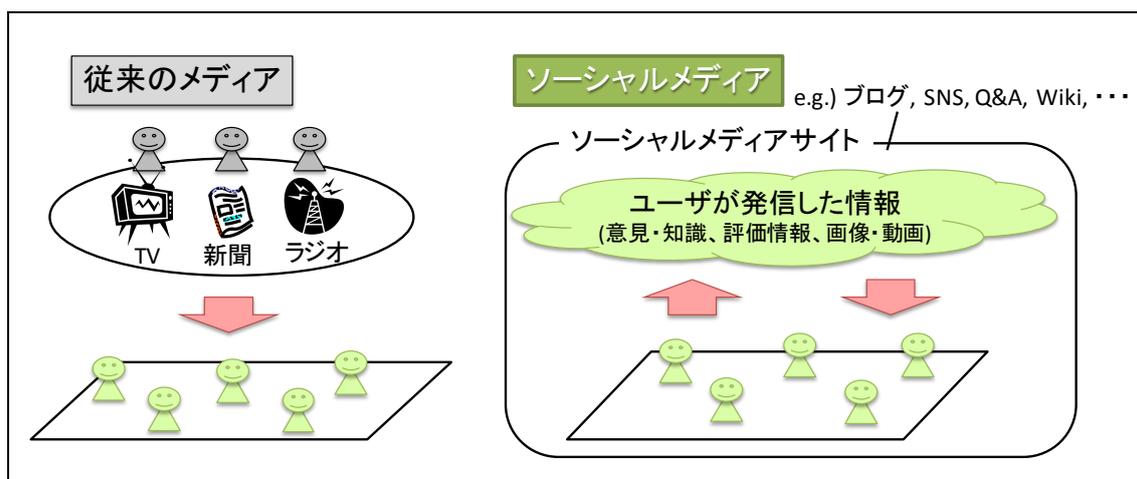


図 1-1：従来のメディアとソーシャルメディアでの情報の生成・提供の違い

1.1.2. ユーザの有用な情報取得を支援する情報共有サイト

ソーシャルメディアを実現する Web サイト（以下、ソーシャルメディアサイト）では、個々のユーザが発信する情報はサイト上に蓄積され、これらの情報は多数のユーザによって参照される情報となる。このような、ソーシャルメディアサイトを介した情報の共有は、従来のメディアでは得られなかった個々のユーザが持つ多様な知識や意見の有用な情報を参加するユーザにもたらしめている。

現在、ソーシャルメディアと呼ばれるサイトは多様に存在しているが、ユーザが持つ多様な知識や意見の共有に特化したサイトは「情報共有サイト」とよばれ、ユーザが有用な情報を効率的に取得できるサイトとしてあげられている [総務省 10a]。総務省のソーシャルメディアサイトの利用実態に関するウェブ調査では、ユーザのソーシャルメディアの利用は、「情報取得」と「コミュニケーション」を目的とした利用が示されており、それぞれ、ユーザの効用が高いサイトとして、「情報取得」に関しては情報共有サイト、「コミュニケーション」に関しては、ブログ、マイクロブログ、SNS、があげられている [総務省 10a]。

情報共有サイトは、参加するユーザの情報共有に特化したサイトであり、ブログや SNS といったユーザのコミュニケーションによる効用も含んだサイトに比べ、個々のユーザが自身の興味関心に基づいて有用な情報を効率的に取得する点において特徴的なサイトである。このような情報共有サイトの例として、価格.com や@コスメなどの商品レビューサイト、あるいは、質問と回答という形式で情報を蓄積しこれを共有する Q&A サイト、また、ブックマークや Web サイト・商品などの評価情報を共有するソーシャルブックマークサイトなどがその代表例としてあげられる。

情報共有サイトにおいてユーザが発信し共有される情報は、商品購入の意思決定や（商品レビューサイト）、Web サイトでは得られないユーザの持つ知識の獲得（Q&A サイト）、膨大な Web サイトや商品などのコンテンツからの有用なコンテンツの選定（ソーシャルブックマーク）など、従来のメディアでは得られなかった個々のユーザが持つ多様な知識や意見の有用な情報の共有を可能にしている。

1.1.3. 情報共有サイトの活性化

情報共有サイトが、参加するユーザに対して価値ある情報を提供し続けることが可能なサイトであるためには、多数のユーザの参加といったサイトの活性化が重要になる。

情報共有サイトは、個人を主体とした情報発信に基づいて情報生成がなされるため、多数のユーザが参加するほどその価値が高まるといったネットワーク外部性を持っている。例えば、情報共有サイトにおいて、ユーザの情報提供の書込みや閲覧、あるいは評価や購買といった行動履歴などの情報受発信行動の情報は、サイト上に蓄積され、これを参照する他のユーザにとって有用な情報源となっている。これらは、多数のユーザが参加することで有用な情報が多く集まる可能性が高まり、ユーザひとりひとりの享受できる利便性が向上しサイトとしての価値を高めている。

ユーザにとって価値ある情報共有サイトを実現するためには、ネットワーク外部性の観点から多数のユーザ参加によるサイトの活性化が重要であり、そのためには、情報共有サイトに参加するユーザのモチベーションである「ユーザの有用な情報取得」を満足させるためのサイトの設計・運用が必要不可欠である。

1.1.4. 情報共有サイトを支える「情報蓄積」と「情報加工」のシステム

情報共有サイトにおいて、ユーザが発信する価値ある情報を「蓄積」し、その蓄積さ

れた情報を有用なかたちに「加工」して提供するためのシステムは、ユーザの有用な情報取得を支援する重要なシステムである（図 1-2）。

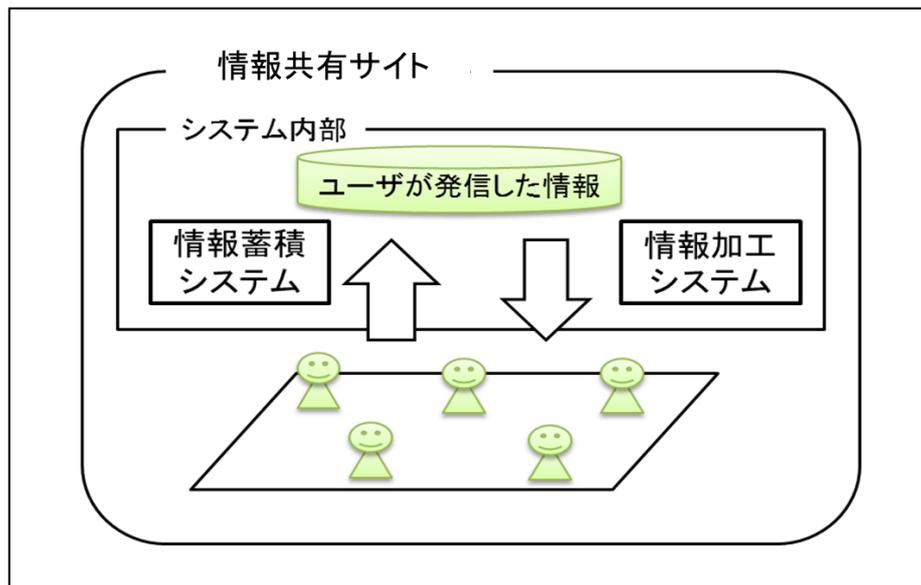


図 1-2：情報共有サイトにおける「情報蓄積」と「情報加工」のシステム

情報蓄積システム

情報共有サイトにおいて、ユーザが発信した情報はシステム内部に蓄積され、この蓄積された情報は、他のユーザにおける新たな情報源として利用される。このため、ユーザが発信する価値ある情報をいかに蓄積するかといった情報蓄積システムは、情報共有サイトが提供可能な情報の価値を左右する重要な役割を持っている。

このような、ユーザが発信する情報を効果的に蓄積するためのシステムは様々に存在する。例えば、その一つとして情報発信におけるユーザの障壁を下げるための情報発信支援のシステムがあげられる。これは、Web 上のフォームに直接入力することで定型化された情報発信を可能にするツール・機能などである。従来、個人は HTML などの専門的な技術を知らなければ自身の Web ページを作成することはできず、意見発信を行うことが可能な人は特定の人に限られていた。しかし、定型化された情報発信を可能とするシステムによって、技術的な知識がない人でも手軽に文章を書き込むことができるようになり、これにより多数の多様な人が発信する情報の蓄積が可能になっている。また、別のシステムの例としては、レビューサイトにおける他者のコメントに対する返信や、コンテンツに対する投票が可能なシステムなどがあげられる。このようなシステ

ムは、ユーザの気軽な情報発信を促し、多数のユーザ発信情報の蓄積を可能にしている。さらに、上記にあげたようなユーザの明示的な情報発信に基づいた情報蓄積のシステムだけでなく、ユーザの商品購買履歴や Web ページの閲覧履歴など、ユーザの暗黙的な情報を記録するような機能も、ユーザの発信する情報を効果的に蓄積するシステムとしてあげられる。

情報加工システム

情報共有サイトにおいて、ユーザの情報発信によって蓄積される情報は膨大、かつその内容は玉石混合である。そこで、ユーザが必要とする情報を提供できるようにするためには、蓄積された情報のなかから価値ある情報を届けるための情報加工のシステムが必要不可欠になる。

このような、蓄積された膨大な情報のなかから、ユーザにとって価値ある情報を加工して提供するためのシステムとして、情報の分類／検索はその主要なシステムとしてあげられる [O'Reilly 05, 梅田 06, 伊藤 07]。情報・コンテンツの分類／フィルタリングに関して、最も一般的な方法としては、文章・画像・動画といった情報・コンテンツ自体の内容を解析する方法があげられる。しかし、このような方法は、膨大で常に変更される情報に対して逐次的な解析を必要し、そのコストは膨大になることから Web 上の情報の分類／フィルタリングの技術としては限界がある。

このようななか、多数のユーザが発信する投票や評価などの他者意見を利用した情報のフィルタリングが有用な技術として注目されている [Surowiecki 06, 大向 07]。その代表的な例は、Google の検索エンジンで利用される PageRank [Lawrence 99] のような検索技術である。これは、ユーザが行う Web ページへのリンク行為を有用なコンテンツへの投票とみなし、このリンク関係によって有用な情報・コンテンツをランキング化するという技術である。ソーシャルメディアサイトにおいても、このような他者意見を利用した情報・コンテンツのフィルタリングの技術は、価値ある情報・コンテンツの発見という観点で多数の Web サイトで適用されている [Schafer 01, Sarwar 01]。

1.1.5. 情報共有サイトにおけるユーザの有用な情報取得の課題

情報共有サイトが、ユーザにとって有用な情報取得が可能なサイトであるためには、「情報提供のモチベーションによる情報蓄積の活発化」と、「膨大な情報からの有用な情報の取得」の支援設計が重要な課題となる。

(1) 情報提供のモチベーション支援による情報蓄積の活発化

情報共有サイトにおいて、ユーザが有用な情報を取得できるためには、サイトに蓄積される情報の活発化が重要な課題となる。

ユーザが発信する情報には、文章投稿や投票行動などの明示的な情報発信によるものと、行動履歴などの暗黙的に発信する情報がある。暗黙的に発信される行動履歴は膨大に蓄積されるが、一方で、文章投稿や投票行動などに基づく情報は、情報発信においてユーザ労力を伴うため、必ずしも有用な情報が多く集まらないといった問題がある。ユーザの情報発信を容易にするツールや機能は、このような多数のユーザが発信する情報の蓄積を支援するシステムである。しかし、このようなツールや機能といった環境を用意しても、ユーザの明示的な情報発信（文書投稿等）には労力を伴うため、必ずしも多数のユーザ情報が蓄積されるとは限らないといった問題がある。例えば、企業内での知識共有を支援するツールにおいても、環境はあるが実際には情報を提供する人が少なく、活発な情報共有がなされないなどの問題も指摘されている[日本情報システム・ユーザ協会 01]。このような、活発な情報提供者の不在による問題は、ユーザの明示的な情報発信による情報提供によって支えられる情報共有サイト（Q&A サイト、レビューサイト、Wiki 等）においては、特に重要な問題である。実際、近年ではこれらのサイトにおいて、ユーザの内発的モチベーションのみに依存した情報提供に陰りが見えつつあるとの指摘も報告されている [ダコプル 10]。

このような点をふまえると、ユーザが発信する情報の蓄積においては、ユーザの情報発信を容易にするツールや機能といった環境を準備するだけでは限界があり、価値ある情報をより蓄積するためには、ユーザの積極的な情報提供を促すための支援が必要になる。特に、ユーザの明示的な情報提供によって情報共有が行われるサイト（Q&A サイト、レビューサイト等）において、価値ある情報を蓄積するためには、ユーザが参加する動機を理解、特に情報提供におけるモチベーションを理解し、それを促すための機能や制度といったシステムは必要不可欠と考えられる。近年の研究において、これらの Web サイトに参加するユーザのモチベーションについては徐々に明らかになってきたが [川浦 99, 山下 05, 三浦 06]、効果的なモチベーション支援の方策に関して客観的な知見は明らかになっていない。

ユーザにとって有用な情報取得を支援する情報共有サイトを実現するためにも、ユーザの情報提供を促すモチベーション支援がもたらす効果を理解し、効果的な支援方策を

検討することは重要な課題である。

(2) 膨大な情報からの有用な情報の取得

ユーザが有用な情報を取得できるためには、蓄積された情報から自身にとって有用な情報を取得するための機能が必要となる。

蓄積される情報が膨大になってしまうと、ユーザにとって必要な情報が埋もれてしまい、それを探すための労力がかかってしまう。特に、ユーザの暗黙的な情報提供行動（閲覧履歴、購買履歴、検索履歴、お気に入り履歴等）によって膨大に蓄積される情報を共有するサイト（はてなブックマーク、Amazon.com の他者情報の参照）において、自身にとって有用な情報のみを取得することは困難である。

膨大な情報からの情報取得手法としては、ユーザの Pull 型による情報取得である検索システムがあげられるが、このようなシステムは、ユーザが自身の望む情報が明確な場合においては有用だが、望む情報が明確でない、あるいは具体化されていない場合においては必ずしも有用な情報取得機能とはいえない。

これに対し、ユーザの受動的な情報取得を支援する「推薦システム」は、ユーザごとにパーソナライズされた Push 型の情報提供が可能であり、ユーザの具体化されていない情報ニーズの具体化や発見において効果的である。なかでも、他者意見を用いた情報推薦は、情報やコンテンツ自体の内容解析を必要としないため、その適用範囲が広く、ニュースサイトや EC サイトなどにおいて、個々のユーザにとってパーソナライズされた有用な情報の取得を支援するシステムとして利用されている [Schafer 01, Sarwar 01]。

しかし、従来の推薦システムにおいて、その有用性の評価は予測精度の適切さといった点に基づいて評価されることが多く、このような予測精度のみを重視した推薦は、ユーザの新たな情報の発見を促す観点においては十分でなく、推薦の満足度の向上においては限界があることが指摘されている [Herlocker 04, McNee 06]。近年では、Novelty（ユーザが未知かつ好みのコンテンツの発見）といった観点から推薦の有用性をとらえ、これを向上させる手法が検討されているが [Ziegler05, 清水 08, 村上 09]、利用者の労力の高さや、コンテンツの特性に限定されてしまうなど、適用範囲の広い手法に関しては十分に検討されていないのが現状である。

このような Novelty の高い推薦は、ユーザの既存の嗜好・興味範囲の過度な固定化を防ぎ、推薦への飽きの解消や、興味の開拓を促す重要な要素であり、これは情報共有サイトを利用するユーザの長期的な満足において重要な要素と考えられる。

ユーザにとって有用な情報取得を支援する情報共有サイトを実現するためにも、Novelty を考慮した多様な情報への接触機会を阻害させない情報推薦の機能の設計は重要な課題である。

1.2. 本研究の目的と意義

1.2.1. 本研究の目的

本研究の目的は、情報共有サイトの活性化を行うためのサイト設計に関する知見を明らかにすることである。そこで、本研究では、情報共有サイトにおけるユーザの参加モチベーションである「ユーザの有用な情報取得」を支援することが多数のユーザの継続的な参加を促し、サイトの活性化に繋がるうえで重要な要件ととらえ、これを支援するためのサイトのシステム設計を検討する。

具体的には、情報共有サイトにおけるユーザの有用な情報取得支援の課題として、「(1) 情報提供のモチベーション支援」と、「(2) 有用な情報の取得支援」をあげ、有効な設計方策の知見を明らかにする。

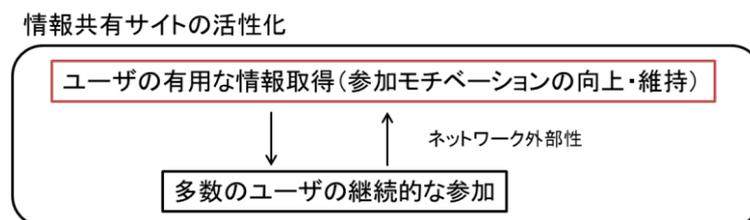


図 1-3 : 情報共有サイトの活性化

1.2.2. 本研究の意義

本研究の意義は、ユーザにとって有用な情報取得を支援する情報共有サイトを実現するための機能や制度に関して、有用な設計・運用方針の知見を示す点である。

現在、ソーシャルメディアサイトは、多様に存在し、それらは実際に多数のユーザのもと運用されている。これらの多くは、ユーザの情報受発信を活発化させるために複数の機能や制度を導入しているが、どのような機能・制度がどう有効であったかという分析は、多数の機能・制度を導入している環境や、ユーザの多様なモチベーションのもとでは困難である。

本研究では、情報共有サイトを支えるシステムと課題を整理し、それぞれに対して有

効な設計・運用アプローチを提案する。本論文で提案する設計・運用方針の知見は、従来の情報共有サイトにおける新たな機能や制度の導入における有効な知見のひとつとして活用できると考える。

1.3. 本論文の構成

本論文の構成を以下に示す。2章では、ユーザに有用な情報取得を可能とする情報共有サイトを実現するうえで課題となる、(1) 情報提供のモチベーション支援と、(2) 有用な情報の取得支援に関して、効果的な設計を行うための課題について関連研究をふまえて整理する。これにより、3章、4章で取り組む研究課題の位置づけを示す。

3章では、(1) 情報提供のモチベーション支援の課題に関して、ユーザの明示的な情報提供に支えられるサイトの代表である Q&A サイトをとりあげる。ここでは、サイト運営者がコントロール可能である外的報酬に焦点をあて、効果的な報酬制度を議論するためのエージェントベースモデルを提案し、外的報酬制度が Q&A サイトにおいて蓄積される知識にどのような影響をもたらすかをエージェントシミュレーションによって分析する [Ogawa 10]。

4章では、(2) 有用な情報の取得支援の課題に関して、サイトに蓄積される個々のユーザの評価行動履歴を加工することによって実現される情報推薦手法の設計課題について取り組む。具体的には、推薦の有用性として Novelty の高い推薦（未知かつ好みのコンテンツの推薦）に着目し、これを利用者の労力が少なく、かつコンテンツの適用範囲が広い推薦手法を提案し、その有用性を被験者実験によって明らかにする [Ogawa 08, 小川 09b]。

5章では、3章と4章の結果をふまえた考察、6章では結論と今後の課題を述べる。図 1-4 は、本研究の構造を示したものである。

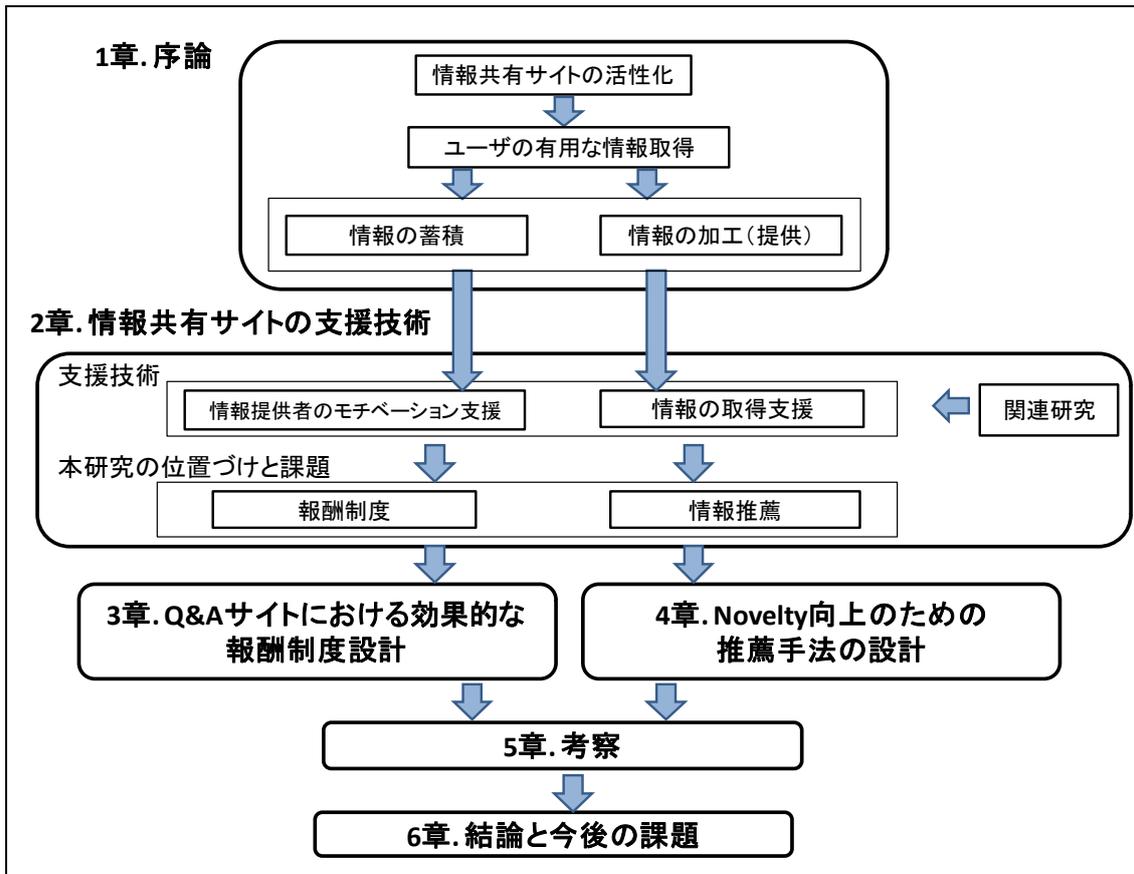


図 1-4 : 本研究の構造

<主要参考文献>

第 3 章

Yuki Ogawa, Hirohiko Suwa, Hitoshi Yamamoto, Isamu Okada, Toshizumi Ohta.: Agent-Based Model of Q&A Community for Effective Pecuniary Payback System, Proc. of the 3rd World Congress on Social Simulation (WCSS2010)

第 4 章

小川祐樹, 諏訪博彦, 山本仁志, 岡田勇, 太田敏澄 : 動的な動的なトピック分類に基づく Novelty を考慮した推薦アルゴリズムの提案, 情報処理学会論文誌 第 50 卷 6 号, pp.1636-1648 (2009)

Yuki Ogawa, Hitoshi Yamamoto, Isamu Okada, Hirohiko Suwa, Toshizumi Ohta.: Development of Recommender Systems Using User Preference Tendencies: An Algorithm for Diversifying Recommendation, Proc. of the 8th IFIP Conference on e-Business, e-Services, and e-Society (I3E 2008), pp.61-73 (2008)

2章. 情報共有サイトの支援技術

本章では、情報共有サイトにおけるユーザの有用な情報取得において課題となる「(1) 情報提供のモチベーション支援」と、「(2) 有用な情報の取得支援」に関して、これを支援する技術を整理し、そこでの課題を述べ、本研究で取り組む位置づけを示す。

2.1. 情報共有サイトにおける情報提供のモチベーション支援技術

本節では、情報提供のモチベーション支援技術を整理する。また、本研究で焦点をあてる外発的モチベーションとしての報酬制度設計について、これを扱うことの必要性を述べ、その設計について従来研究の課題を述べる。

2.1.1. 情報提供のモチベーション

ユーザが有用な情報を取得できるためには、その情報源の活発化が必要である。このためには、ユーザの情報提供を促すためのモチベーション支援が重要な方策となる。

情報共有サイトにおいて、ユーザが発信する情報には、文章投稿や投票行動などの明示的な情報発信によるものと、行動履歴などの暗黙的に発信する情報がある。暗黙的に発信される行動履歴は膨大に蓄積されるが、一方で、文章投稿や投票行動などに基づく情報は、情報発信においてユーザ労力を伴うため、必ずしも有用な情報が多く集まらないといった問題がある。情報提供行動に伴う労力を軽減させるためのツールや機能はいくつかあげられるが、このような物理的な環境は整っていても、情報提供のモチベーションがなければ活発な情報共有がなされないなどの問題も指摘されている[日本情報システム・ユーザ協会 01]。特に、ユーザの明示的な情報提供によって情報共有が行われるサイト (Q&A サイト、レビューサイト、Wiki 等) において、モチベーション支援は必要不可欠と考えられる。

ユーザの情報提供のモチベーションを向上させるための方法を考えたとき、その方法は、外発的モチベーションと内発的モチベーションとに分けられる。外発的モチベーションとは、行動に対する外的報酬によって高められるモチベーションである。ここで、報酬には、金銭やポイントなどの物理的な報酬や、他者からの承認や評価といった非金銭的なものがあげられる。一方、内発的モチベーションとは、外的報酬に基づかないモチベーションであり、ユーザの自発的な意思によって行動し満足感を得るモチベ

ンである。

2.1.2. 情報提供の外発的モチベーション

情報共有サイトにおけるユーザの情報提供のモチベーションを促すうえで、外発的・内発的モチベーションのどちらに対する支援も重要な要素である。特に、多様な興味関心を持ったユーザが参加する情報共有サイトにおいては、多数のユーザ参加によってサイトの活性化を促すためには、内発的モチベーションだけでなく外発的モチベーションに対する支援も重要な要素になってくると考える。

情報共有サイトにおける情報提供者のモチベーションはサイトによってさまざまなものがある。例えば、Wikipedia においては、参加するユーザの動機は楽しみやイデオロギーといった内発的モチベーションが最も重要な動機であることが指摘している [Oded 07]。また、Q&A サイトにおいては、回答を提供するユーザの動機は、質問者を助けたいという援助的動機といった内発的モチベーションだけでなく、他者からの評価や報酬などの報酬的動機といった外発的モチベーションも、回答提供のモチベーションとなっていることが調査結果によって明らかになっている [三浦 08]。

このような情報提供のモチベーションを支援する方法として、内発的モチベーション支援に関し、[青山 08] らは、企業内のユーザ参加型の知識共有システムの活発化のために、ユーザに対して組織方針と自身の目的を認識・共有することでユーザの内発的モチベーションが高まることを示している。このような内発的モチベーションは、ユーザ自身の知的好奇心や、自発的に思考して問題を解決するといった行動に付随するモチベーションであり、ユーザの継続的な行動を促すモチベーションとして重要な要素である。しかし、近年、このようなユーザの内発的なモチベーションによって情報共有がなされる Wikipedia において、情報提供に陰りが見えつつあるとの指摘も報告されている [ダコプル 10]。

情報提供の活発化のために、ユーザの内的モチベーションを促すための目的を設定し、これを共有させるといった支援策も考えられるが、このような方法では対象とするユーザが限定されてしまい、サイトの設立初期などのユーザが少ない場合においては必ずしも効果的とは限らないと考えられる。特に、多様な興味関心を持ったユーザが参加する Web 上での情報共有サイトにおいては、多数のユーザ参加によってサイトの活性化を促すためには、内発的モチベーションだけでなく外発的モチベーションに対する支援も重要な要素になってくる。このような外的報酬によるモチベーション支援は、サイトの運営者がその大きさを操作可能な支援であり、多様なユーザが参加する Q&A サイトやレ

ビューサイトにおいて、金銭、ポイント、他者評価などによる報酬によって情報提供のモチベーションを促すための制度が取り入れられている。

2.1.3. 情報共有サイトにおける報酬制度の重要性

報酬制度といった情報提供者への外的報酬によるモチベーション支援は、情報共有サイトへの有用な情報の蓄積のために、サイトの設計・運用者が関与可能な重要な支援策である。

情報共有サイトにおいてユーザに提供される情報は、他のユーザが発信した情報である。このため、サイトに参加するユーザが有用な情報を取得するためには、情報提供を行うユーザの情報発信行動をいかに促すかが重要な課題となる。

個々のユーザの情報発信を促す方策として、情報発信の負担を下げるためのツールの整備といったものもあげられるが、これらのツールの利用はユーザの自発的な行動に依存してしまうため、必ずしも効果的な情報蓄積になるとは限らない。例えば、Q&A サイトや、レビューサイト等においては、情報の蓄積が、文章投稿といったユーザのコストの伴う明示的情報発信によって行われるため、情報自体がサイトに蓄積されないといった問題も生じる。また、蓄積される情報においても、ユーザの自発的な情報提供に基づいているため、必ずしも他のサイト参加者にとって有用な情報が蓄積されるとは限らない。これは、例えば、Q&A サイトにおける質問者の意図した回答が集まらないといった情報の質の問題や、Amazon などのレビューサイトにおける評価されるコンテンツの偏りといった問題によって、サイト全体として有用な情報が蓄積されないといった問題が考えられる。

このようななか、報酬制度といった情報提供者へのモチベーション支援策は、情報共有サイトへの有用な情報の蓄積のために、サイトの設計・運用者が関与可能な支援策である。例えば、Q&A サイトにおける報酬制度は、情報提供者である回答者に外的報酬を与えることで、回答者のモチベーションの向上につながり、質問者が望む情報の取得を支援することが可能になると考えられる。また、商品などを評価するレビューサイトにおいては、評価情報の無い・少ないコンテンツに対して、報酬といったインセンティブによって情報提供者のレビュー行動を促すことで、多様な情報を収集し、サイト参加者にとって有用な情報を提供するといったことが可能になると考えられる。

このようなサイトの設計・運用者が情報提供者のモチベーションに関与可能な報酬制度に関して、効果的な制度設計を検討することは、情報共有サイトへの有用な情報蓄積を促すうえで重要な課題になると考えられる。

2.1.4. Q&A サイトにおける報酬制度

情報提供者への報酬の付与に関しては、サイトからの報酬（例：レシピ投稿者への金銭報酬「nanapi ワークス¹」）や、ユーザ間による報酬（例：質問者から回答者への報酬「Q&A サイト」）に分けられる。しかし、多数のユーザが参加し望む情報も参加ユーザによって異なるような情報共有サイトにおいては、サイトによる報酬では、サイト運営者という特定の個人が、有用な情報というものを定義・判断することは困難であるため、ユーザ間による報酬といった情報を受け取るユーザの評価が行われる方法がより重要な方法となると考えられる。

このようなユーザ間による報酬を用いているサイトとして、Q&A サイトがその代表例としてあげられる。Q&A サイトでは、情報を必要とする質問ユーザが明示化されることで、回答者の回答行動を促し、また、質問を受け取るユーザからの評価によって、個々のユーザにとってより有用な情報が蓄積される。このような、質問・回答による情報共有は、質問者が文章というかたちで疑問を投稿することができ、検索システムを使いこなせないユーザの利用や、蓄積されていない情報を他のユーザから収集するというかたちでの情報取得を可能としている。近年、このような Q&A の仕組みは、Twitter 上や SNS 上での Q&A サービスの導入など²³、多数のサイトで取り入れられてきていることから、ユーザの有用な情報取得を支援するための重要な機能となりうると考えられる。

2.2. 情報共有サイトにおける有用な情報の取得支援技術

本節では、情報共有サイトにおける個々のユーザの情報選択を支援する技術を整理し、本研究で焦点をあてる情報推薦に関してその課題を述べる

2.2.1. 情報提供の形態

情報提供を行うサービスやサイトにおいて、その情報提供の形態は、Pull 型と Push 型に分類される [石川 06]。

¹ nanapi ワークス : <https://works.nanapi.jp/>

² Q&A なう : <http://qa-now.com/>

³ Facebook Question : <http://www.facebook.com/questions/>

Pull 型情報提供

Pull 型情報提供とは、情報を取得するユーザの情報検索・入手によって情報が提供される形態のことである。このような情報提供を備えるサイトにおいては、ユーザは能動的な検索システムの利用によって、自身の望む情報の取得が可能である。

Push 型情報提供

Push 型情報提供とは、情報提供を行うサイトやシステムが、ユーザに対して積極的に情報をフィルタリングして提供する形態である。これは、例えば、サイトの管理者が登録したルールに基づいてユーザに情報提供を行うものや、ユーザが登録した嗜好情報に基づいて情報提供を行うもの、多数のユーザの行動履歴から個々のユーザの嗜好を推測し情報提供を行うものなど、その実現方法は多様に存在する[石川 06, 神 07]。

2.2.2. 情報共有サイトにおける情報推薦システムの重要性

情報共有サイトにおいて、検索システムや推薦システムは、蓄積された膨大な情報からユーザが有用な情報を選択するうえで重要なシステムである。なかでも、情報推薦システムは、膨大な情報のなかから、自身の嗜好に合った情報のみを受動的に取得することが可能な有用な機能である。

検索システムの利用はサイトに蓄積された情報からの情報取得方法として、一般的な方法である。検索システムは、ユーザは自身の望む情報を検索クエリというかたちで要求し、それに対する結果提示の繰り返しによって、自身が望む情報を取得する方法である。しかし、ユーザが効率的な情報収集を行うためには、検索クエリを与えるユーザのスキルを要するため、すべてのユーザが有効な情報取得を行える手法とは限らないといった問題がある。また、このようなシステムは、ユーザが自身の望む情報が明確な場合においては有用だが、望む情報が明確でない、あるいは具体化されてない場合においては必ずしも適した情報取得機能とはいえない。

一方、情報推薦システムは、ユーザが自身の嗜好を表現する情報をシステムに提供することで、自身の望む情報を受動的に取得できるシステムである。ここで、ユーザがシステムに提供する情報は、自身が積極的に登録する情報（デモグラフィック属性やキーワード）だけでなく、商品購買履歴や、サイトの閲覧履歴など、ユーザが暗黙的に発信する行動履歴などの情報も用いられる。このような情報推薦システムによるユーザの情報

取得支援の機能は、膨大な情報のなかから、自身の嗜好に合った情報のみを受動的に取得することが可能な有用な機能である。

さらに、推薦システムは、ユーザの情報取得の支援だけでなく、ユーザの情報提供行動を促すといった観点においても重要な機能と考えられる。推薦による Push 型の情報提供は、ユーザの興味関心を喚起させることで新たな利用行動を促し、ユーザの暗黙的な情報発信のデータ（閲覧履歴、購買履歴等）の蓄積といった観点での効果も期待できる。例えば、レビューサイトにおいては、サイトにおいて必要な情報（コンテンツへの評価やレビュー等）を、推薦によって積極的に提示し、ユーザの興味を喚起させることで、ユーザの情報提供行動を促すといった活用、あるいは、Q&A サイトのような膨大な質問が寄せられるサイトにおいては、質問に最も適した回答者に対して質問推薦を行うことで効率的な情報蓄積を支援したり、分野の異なる多数のユーザに対して同一の質問を推薦することで多様な回答を集めたりといった利用によって、サイトへの有用な情報を蓄積したりといった場面での活用も考えられる。

2.2.3. 他者意見を利用した情報推薦システム

情報推薦システムの実現手法は、大きく、情報・コンテンツの内容を利用するものと、他者意見を利用するものに分類される [神畠 07]。

情報・コンテンツの内容を利用するものは、情報・コンテンツの属性情報から抽出した特徴情報と、ユーザの嗜好を表現した特徴情報を比較し、それらが類似するものを推薦対象としてユーザに提供する手法である [Pazzani 96, Mooney 99]。

一方、他者意見を利用した情報の推薦システムは、ユーザの嗜好情報（主に、行動履歴が用いられる）を共有し、推薦対象ユーザと似た嗜好をもっているユーザの嗜好情報をもとに、ユーザの嗜好を推測し、推薦対象となるものを提供する手法である [Resnick 94]。例えば、情報共有サイトの一つであるソーシャルブックマークサイト^{4,5}では、Web を通じてブラウザのブックマークを共有することで、他者が興味を持っているトピックや、多くの人に興味を持っている情報・コンテンツの発見を実現している。

このような、他者意見に基づいて情報をフィルタリングし推薦するシステムは、協調フィルタリング [Resnick 94, Adomavicius 05, 神畠 07] やソーシャルフィルタリングと呼ばれ [神畠 06]、ユーザの嗜好・評価情報があればどんな情報・コンテンツも対象と

⁴ del.icio.us, <http://del.icio.us/>

⁵ はてなブックマーク, <http://b.hatena.ne.jp/>

なるため、その適用範囲が広く、多数のサイトで取り入れられている [Schafer 01, Sarwar 01]。

2.2.4. 情報共有サイトにおける情報推薦システムの課題

他者意見を利用した推薦システムは、膨大な情報からユーザの有用な情報取得を支援する有用な機能であるが、その有用性の評価は予測精度の適切さといった点に基づいて評価されることが多い。しかし、このような予測精度のみを重視した推薦は、ユーザの新たな情報の発見を促す観点においては十分でなく、推薦の満足度の向上においては限界があることが指摘されている [Herlocker 04, McNee 06]。近年では、Novelty（ユーザが未知かつ好みのコンテンツの発見）といった観点から推薦の有用性をとらえ、これを向上させる手法が検討されているものの [Ziegler05, 清水 08, 村上 09]、利用者の労力の高さや、コンテンツの特性に限定されてしまうなど、適用範囲の広い手法に関しては十分に検討されていないのが現状である。

このような Novelty の高い推薦は、ユーザの既存の嗜好・興味範囲の過度な固定化を防ぎ、推薦への飽きの解消や、興味の開拓を促す重要な要素であり、これは情報共有サイトを利用するユーザの長期的な満足において重要な要素と考えられる。ユーザにとって有用な情報取得を支援する情報共有サイトを実現するためにも、Novelty を考慮した多様な情報への接触機会を阻害させない情報推薦の機能の設計は重要な課題である。

2.3. 本研究の位置づけと課題

本研究では、情報共有サイトにおけるユーザの参加モチベーション（ユーザの有用な情報取得）を支援することが多数のユーザの継続的な参加を促し、サイトの活性化に繋がるうえで重要な要件ととらえ、これを支援するためのサイトの設計・運用に関して検討する。具体的には、情報共有サイトにおけるユーザの有用な情報取得の課題である以下の2つの課題に関し、それぞれ有効な支援策に関する知見を明らかにすることを目的とする。

2.3.1. 情報提供のモチベーション支援

Q&A サイトにおける報酬制度設計の課題・・・(3章)

ユーザが有用な情報を取得できるためには、その情報源の活発化が必要である。ただし、情報提供の労力を軽減するためのツールや機能などだけではユーザの明示的な情報提供（文章投稿等）にもとづく情報の蓄積には限界があり、このためにはユーザの情報提供を促すためのモチベーション支援が重要となる。

情報共有サイトにおけるユーザの情報提供のモチベーションを促すうえで、外発的・内発的モチベーションがあげられる。情報提供の活発化のために、ユーザの内的モチベーションを促すための目的を設定し、これを共有させるといった支援策も考えられるが、このような方法では対象とするユーザが限定されてしまい、サイトの設立初期などのユーザが少ない場合においては必ずしも効果的とは限らないと考えられる。特に、多様な興味関心を持ったユーザが参加する Web 上での情報共有サイトにおいては、多数のユーザ参加によってサイトの活性化を促すためには、内発的モチベーションだけでなく外発的モチベーションに対する支援も必要な要素になってくる。このような外的報酬によるモチベーション支援は、サイトの運営者がその大きさを操作可能な支援であり、多様なユーザが参加する Q&A サイトやレビューサイトにおいて、金銭、ポイント、他者評価などによる報酬によって情報提供のモチベーションを促すための制度が取り入れられている。

また、報酬の与え方に関して、多数のユーザが参加し望む情報も参加ユーザによって異なるような情報共有サイトにおいては、サイトによる報酬では、サイト運営者という特定の個人が、有用な情報というものを定義・判断することは困難であるため、ユーザ間による報酬といった情報を受け取るユーザの評価が行われる方法がより重要な方法となると考えられる。このようなユーザ間による報酬を用いているサイトとして、Q&A

サイトがその代表例としてあげられる。

Q&A サイトでは、情報を必要とする質問ユーザが明示化されることで、回答者の回答行動を促し、また、質問を受け取るユーザからの評価によって、個々のユーザにとってより有用な情報が蓄積される。Q&A サイト参加者の動機には様々なものが考えられるが、運営者が設計可能な動機づけの方法としては、やり取りされた Q&A に対して、利用者相互、もしくは運営者がどのような対価を支払うかといった報酬制度があげられる。この報酬制度について、実際の Q&A サイトにおいては、コミュニティ運営者の方針の違いから異なる取り入れ方がなされている。

しかし、このような報酬制度の違いが、長期的な観点において Q&A サイトでやり取りされる質問、あるいは回答の数や質などの Q&A サイトとしての成果にどのように影響し、結果的にサイトに蓄積される情報にどのような効果をもたらすかといったメカニズムについて、客観的な知見は明らかになっていない。また、Q&A サイトにおける報酬制度を分析、あるいは効果的な設計を議論するための方法モデルについても十分な議論がなされていないといった課題がある。

以上の課題をふまえて、本研究では、Q&A サイトに焦点をあて、報酬制度を議論するためのモデルとして、[三浦 08] の研究によって示された心理的動機をベースとしたエージェントベースモデルを構築する。このようなエージェントベースモデルは、質問者・回答者における効用や選択などの相互作用過程を、さまざまな報酬制度について観察することが可能であるといった利点がある。また、提案したモデルにもとづきシミュレーション実験を行うことで報酬制度の導入効果を分析し、報酬制度の導入方針に関する知見を得ることを目的とする。

このような、ユーザの情報提供を促すモチベーション支援がもたらす効果を理解し、効果的な支援方策を検討することは、ユーザにとって有用な情報取得を支援する情報共有サイトを実現するためにも重要な課題である。

2.3.2. 有用な情報の取得支援

Novelty 向上のための情報推薦手法… (4 章)

情報共有サイトにおいて、検索システムや推薦システムは、蓄積された膨大な情報からユーザが有用な情報を選択するうえで重要なシステムである。なかでも、情報推薦システムは、膨大な情報のなかから、自身の嗜好に合った情報のみを受動的に取得することが可能な有用な機能である。

推薦システムは、膨大な情報からユーザの有用な情報取得を支援する有用な機能であるが、このような予測精度のみを重視した推薦は、ユーザの新たな情報の発見を促す観点においては十分でなく、推薦の満足度の向上においては限界があることが指摘されている[Herlocker 04, McNee 06]。近年では、Novelty（ユーザが未知かつ好みのコンテンツの発見）といった観点から推薦の有用性をとらえ、これを向上させる手法が検討されているものの [Ziegler05, 清水 08, 村上 09]、利用者の労力の高さや、コンテンツの特性に限定されてしまうなど、適用範囲の広い手法に関しては十分に検討されていないのが現状である。

以上の課題をふまえ、本研究では、情報共有サイトにおいて個々のユーザが発信する多数の評価履歴情報を用いて、利用者の労力が少なく、かつコンテンツの適用範囲が広い Novelty 推薦手法を提案し、その有用性を被験者実験によって明らかにする。

このような Novelty の高い推薦は、ユーザの既存の嗜好・興味範囲の過度な固定化を防ぎ、推薦への飽きの解消や、興味の開拓を促す重要な要素であり、ユーザにとって有用な情報取得を支援する情報共有サイトを実現するうえで重要な課題である。

3章. Q&A サイトにおける効果的な報酬制度設計

3.1. 研究背景

近年、インターネット上において、コミュニティ内の参加者同士が互いに質問と回答を投稿しあう Q&A サイトと呼ばれる知識共有コミュニティが数多く出現し、大きな発展を遂げている（Yahoo!知恵袋⁶、OKWave⁷、教えて!goo⁸、はてな人力検索⁹）。

Q&A サイトの運営者・設計者の立場にとって、コミュニティメンバーの活発な Q&A の利用や、そのコミュニティにとって効果的な成果をもたらすための運営方針や制度設計の策定は、サービスの維持、コミュニティの活性化・発展を考えるうえで重要な課題である。

現在では、企業におけるナレッジマネジメントの方策としても知識共有コミュニティの構築は模索されている。コミュニティの運営方針や制度設計の知見を探る方法として、成功した個々のサービスに対する事例ベースでの検証は可能である。しかし、これらのケーススタディのみから、運営方針や制度設計がコミュニティの活性化に対してどのような効果をもたらすかのメカニズムについて探ることは困難である。運用方針や制度設計として、参加者のモチベーションの向上・維持といった観点に関する知見を明らかにすることは、特に重要な課題である。

Q&A サイト参加者の動機には様々なものが考えられるが、運営者が設計可能な動機づけの方法としては、やり取りされた Q&A に対して、利用者相互、もしくは運営者がどのような対価を支払うかといった報酬制度があげられる。この報酬制度について、実際の Q&A サイトにおいては、コミュニティ運営者の方針の違いから異なる取り入れ方がなされている。例えば、「Yahoo!知恵袋」では、お礼の言葉やベストアンサーといった非金銭的な報酬制度を用いているが、「人力検索はてな」では、質問者が回答の質に応じて「はてなポイント」と呼ばれる換金性のあるポイントを支払う報酬制度を用いている。これらの報酬制度の導入方針は、非金銭的な報酬制度の場合では質問者の返報にか

⁶ Yahoo!知恵袋 : <http://chiebukuro.yahoo.co.jp/>

⁷ OKWave : <http://okwave.jp/>

⁸ 教えて!goo : <http://oshiete.goo.ne.jp/>

⁹ はてな人力検索 : <http://q.hatena.ne.jp/>

かるコストを下げ、逆に、金銭的な報酬制度の場合では回答者の回答モチベーションを上げることを期待した運用方針と考えられる。

しかし、このような報酬制度の違いが、長期的な観点において Q&A サイトでやり取りされる質問、あるいは回答の数や質などの Q&A サイトとしての成果にどのように影響を与えているかといったメカニズムについて、客観的な知見は明らかになっていない。さらに、制度の導入効果を分析する方法についても、個々の利用者の特性の違いや、複数の制度を導入しているといった現実の Q&A サイトの複雑さを考えると、行動データのみを主体とした回帰分析的な手法では限界があると考えられる。

本稿では、Q&A サイトにおける利用者間の知識取引（質問・行動・返報）をエージェントベースでモデル化し、シミュレーション実験を行うことで報酬制度の導入効果を明らかにする。具体的には、コミュニティが制度として設定する返報における報酬の高さが、回答の数、回答の質、質問者の質問のしやすさといった観点においてどのような効果をもたらすのかを明らかにする。また、質問者と回答者の知識傾向の同質性が高い、あるいは低いといった異なるコミュニティ環境に対して、シミュレーションによるシナリオ分析を行うことで、コミュニティ環境にとって有効な報酬制度の導入方針に関する知見を得ることを目的とする。

本章の構成を以下に示す。3.2 節において、関連研究として知識共有コミュニティに関する研究を述べ、本研究の位置づけを示す。3.3 節において、提案する Q&A サイトのエージェントベースモデルを述べ、3.4 節において報酬制度の導入効果、コミュニティ環境による報酬の効果の違いについてシミュレーション実験を行い、3.5 節において考察、3.6 節でまとめと今後の課題について述べる。

3.2. 先行研究と本研究の位置づけ

本節では、Q&A サイトの参加動機に関する研究と、エージェントベースシミュレーションに関する研究を述べ、本研究の目的と位置づけを示す。

3.2.1. Q&A サイトに関する研究

[三浦 08] は、Q&A サイトにおける利用者の参加動機を明らかにすることを目的とし、Yahoo!知恵袋を対象に、利用者に対する質問紙調査を実施している。その結果、回答投稿の参加動機として、(1)援助的動機、(2)互酬的動機、(3)社会的動機、(4)報酬的動機の4つの因子を明らかにしている。(1)援助的動機は質問者を助けることに、(2)互酬的動機

は回答投稿による将来的なメリット獲得への期待や過去の経験に対する返報に、(3)社会的動機は投稿という行為そのものの社会的意味に、(4)報酬的動機は回答投稿によって得られる(即時的な)メリットに、それぞれ重きを置く傾向を表している。比較分析により、Q&A サイトは基本的には利用者の積極的なサポートによって成り立っているが、同時にその他の参加動機として、回答者の回答行動に伴う自身への直接的な報酬も参加のインセンティブの一つとして影響することが示されている。この分析結果より、Q&A サイトにおける利用者の参加動機は、利用者の援助的特性といった静的な要因だけでなく、質問と回答の知識取引といった動的に変化する相互作用の結果によっても影響を受けるものとして捉える必要があるといえるだろう。

コミュニティの活性化や発展に対する研究の分析のアプローチに関して、コミュニティのサイズと活動状況の因果関係を分析した研究[Koch 04]や、多数の OSS 開発コミュニティのダイナミクスを測定してコミュニティの発展プロセスを数学的にモデル化する研究[Wu 07]などがあげられるが、その多くが、全体的な挙動を表現する数式に基づいてコミュニティの挙動を表現するといったトップダウン的なアプローチによるものである。しかし、多様な利用者が相互作用し、その連鎖によって知識が生成される Q&A サイトに対し、このようなトップダウン的なアプローチでは複雑なメカニズムを記述することは困難であり、モデル化にも限界がある。

最近では、このようなトップダウン的なアプローチのほかに、コミュニティ参加者の行動パターンをエージェントとしてモデル化し、エージェントシミュレーションによってコミュニティの発展プロセスをモデル化しようとする試みもある[鳥海 08、山田 08]。

3.2.2. エージェントベースシミュレーションによるに関する研究

エージェントベースシミュレーションと呼ばれるアプローチは、社会における個々の行為者をエージェントとしてモデル化し、さまざまな条件のもとでシミュレーションを行って実行結果を分析することによって新たな知見を得ることをねらった分析手法である(エージェントベース社会シミュレーションとも呼ばれる(以降、ABSS))。現在では、テロ対策、感染症対策、証券取引、避難誘導など個々人に多様な人間の意思決定が中心となる問題への有力なアプローチとして、多様な分野で用いられてきている[7][8]。この ABSS については、シミュレーションの課題設定やモデルの設計のプロセスなど、その方法論に関してはさまざまな議論がなされている[Gilbert 99][寺野 01][大堀 08]。

Gilbert ら[Gilbert 99] は、ABSS におけるモデルクラスを、Abstract Models、Middle Range Models、Facsimile Models といった3つに分類しており、それぞれ、Abstract Models では、事象表現はできないが基礎的な社会プロセスを示すようなモデル（例：分居モデル[Schelling 78]）、Middle Range Models では、現実市場における事象と質的な類似度を満たすモデル（例：評判モデル[山本 03]、規範モデル[倉橋 01]）、Facsimile Models では現実の特定の市場を正確に再現するモデル（例：感染症対策モデル[金谷 08]）といった異なるモデルとして ABSS 研究を位置付けている。また、大堀ら[大堀 08] は、さらにこれら3つのモデルを目的ベースで分類することで ABSS 研究の意義を整理している。例えば、Abstract Models では、理論的な分析を目的とし、社会現象に対する理論提供の可否が適し、Middle Range Models では、シナリオ分析を目的とし、政策評価への利用可能性の議論、さらに、Facsimile Models では、将来予測を目的とし、現実社会におけるデータとの整合性に関する議論が適していると述べている。

また、Web サイト上のコミュニティのモデル化を行った研究として [山田 08] らの研究がある。[山田 08] らは、OSS 開発コミュニティや Q&A サイトを含めて知識コミュニティとし、抽象化したモデルの構築を行っている。しかし、目的や成功の種類の異なるコミュニティについて [Lin 07]、それぞれ特有の事象を理解するという点においては限界があるといえる。

3.2.3. 本研究の位置づけ

本研究では、Q&A サイトに焦点をあて、報酬制度を議論するためのエージェントベースモデルを提案する。このようなエージェントベースモデルは、質問者・回答者における効用や選択などの相互作用過程を、さまざまな報酬制度について観察することが可能であるといった利点がある。また、提案するモデルのモデル化に関しては、[三浦 08] の研究によって示された心理的動機をベースとしたエージェントベースモデルを構築し、シミュレーション実験を行うことで報酬制度の導入効果を分析する。これは、ABSS 研究のなかでは Middle Range Models といった研究に位置付けることができる。

本研究では、提案したモデルをもとにシミュレーションによるシナリオ分析を行うことで、有効な報酬制度の導入方針に関する知見を得ることを目的とする。

3.3. Q&A サイトのエージェントベースモデル

本節では、利用者間の質問・回答・返報といった相互作用（以下、知識取引）、利用者（質問者、回答者）、報酬制度、利用者の効用、コミュニティ環境（質問者と回答者の同質性）のモデル化について述べる。

3.3.1. 知識取引のモデル化

我々は、Q&A サイトにおける質問者と回答者の相互作用を、(1)質問者の質問投稿、(2)回答者の質問への回答投稿、(3)質問者の回答者への返報投稿といった知識取引のプロセスとしてモデル化する。

さらに、各プロセスで質問者や回答者が得られる利得・コストを変数化し、質問者と回答者それぞれの効用を定式化する。なお、このモデル上での利得は、[三浦 08]での利用動機を満たすものを意味するものとして用いている。

シミュレーションのモデル設計におけるモデルの抽象度に関しては、シミュレーションを用いる目的によってことなる設計方針が必要である。[Gilbert 99]は、将来予測を目的とした場合には詳細なモデル化が必要であるが、現象の理解という目的においては、本質的な要素のみのシンプルなモデル化が重要であると述べている。本研究は、報酬制度の違いがコミュニティにもたらす効果を明らかにすることを目的としているため、報酬が影響を与える動機について焦点をあてたモデル化を行うこととし、[三浦 08]であげた回答者の動機因子の社会的動機と互酬的動機についてはモデルには組み込まないものとした。これら二つの因子を除いた理由として、社会的動機については、Q&A サイトは SNS などのコミュニケーションを主体としたコミュニティではない点と、[三浦 08]での分析結果において他の因子よりも影響の小さい因子であったこと、また、互酬的動機についても、報酬の効果とは直接的に関与しない要素である点からモデル化の対象外とした。

以下の項では、質問者と回答者における知識取引のプロセス（(1)質問投稿、(2)回答投稿、(2)返報投稿）について詳細を述べる。

(1) 質問者の質問投稿

質問者は、自分の疑問にもとづいて質問をコミュニティに投稿する。このとき、質問者の行動として、質問を詳細に記述すればするほど、よりマッチした回答（質問者の意

図した回答)が集まり高い効用を得られると考えられる。しかし、質問自体は質問者が記述する文章という形で表現されるため、質問投稿にはいくらかの質問作成のための労力を有するものと考えられる。また、どれだけ真面目に文章を記述するかは、質問者ごとに異なっていると考えられる。以上より、質問者の質問投稿は、質問者が自身の疑問と、努力特性によって生成される質問をコミュニティに投稿するものとする。

(2) 回答者の質問への回答投稿

回答者は、投稿された質問に対し回答を行うことで満足やポイント等の利得を得る。このとき、回答者の行動としては、自分の知識で応えられる質問、あるいは、お礼（以下、返報）をしてくれそうな質問者かどうかを考慮して質問を選択すると考えられる。また、このとき回答する文章も、どれだけ質問者の意図に合った回答を記述するかは回答者の回答努力によって異なっていると考えられる。以上より、回答者は、自身の知識特性と、質問者の返報特性を考慮して質問を選択し、自身の努力特性にもとづいて質問に合った回答を投稿するものとする。

(3) 質問者の回答者への返報投稿

質問者は、質問に回答があった場合に、回答者に対し返報を行う。しかし、集まる回答は必ずしも質問者にとって意図した回答でない場合もあるため、自分の質問意図（疑問）にある程度マッチした回答に対してのみ返報を行うものとする。

3.3.2. 利用者のモデル化

本節では、3.1 節の 1~3 項であげた知識取引のプロセスにもとづき、Q&A 質問者・回答者と、そこで投稿される質問・回答をモデル化する。

なお、本研究で扱う知識とは、個人の持つ知識特性を抽象的に表現したものである。個人の特性についてのモデル化に関して、[Axelrod 97]では、文化の流布のシミュレーションモデルにおいて、ある個人の文化特性を抽象的に記述できるようにするために、文化特性をさまざまな次元（言語、宗教、技術、衣服のスタイルなど）を表す数字のリストによってモデル化している。また、[鳥山 09]では、組織内での個人が持つスキーマ（ものの見方、判断基準）が組織内でどのように伝搬されていくかをシミュレーションモデルで分析するために、個人の持つスキーマを数字列のリストで表現している。本

研究においても、これらのモデル化をふまえ、個人の持つ知識特性と個人間で取引される質問・回答について、その特性を抽象化して表現したものとしてモデル化する。具体的には、ジャンルなどのある特定の知識単位に対する知識特性を 0、1 で表現し、これらのまとまりをビット列として表現する。なお、現実的には、ある知識単位の特性は多次元であることが考えられるが、本研究では簡素化のために 0、1 といった二次元の値で抽象化したモデル化とする。

また、参加者の役割については、本稿では互酬性を扱わないことと、モデルの簡略化のために、質問者・回答者の役割は固定的なものとしてモデル化した。現実的には質問者・回答者の役割は固定ではなく、その時々によって変わりうるものが考えられる。本研究では、質問者と回答者の相互作用による Q&A サイトの挙動を観察する。そのために、同一エージェントが両者の役割を持つと、援助性や報酬制度の導入効果の分析が困難になるため、本研究では役割固定のモデル化とした。また、[三浦 08]においても、参加者の利用行動のうち全体の約半数は質問のみ、あるいは回答のみの固定的な利用が行われていることから、役割固定のモデル化妥当であると考えられる。役割が動的に変化するモデルについては今後の検討課題とする。

以下の項では、質問者のモデル、回答者のモデルの詳細について述べる。

質問者のモデル

疑問 ($Prop^Q$)

本稿では、質問を投稿する質問者 i の疑問（質問者の質問の特性） $PropQ_i$ を以下のように表現する。ここで、各ビットは、ある知識単位における質問者 i の知識特性を表わしている（例えば、音楽という知識単位に対する特性[邦楽、洋楽]、スポーツという知識単位に対する特性[野球、サッカー]など）。

$$prop_i^Q = [x_1, x_2, \dots, x_k] = [x_n \in \{0,1\}] \quad (\text{式 1})$$

質問努力 ($Effort^Q$)

投稿される質問の内容は、質問者が記述する文章という形で表現されるため、質問がどれだけ詳細・曖昧に記述されるかは、質問者がどれだけ労力をかけて文章を作成するかといった質問者の質問努力に依存すると考えられる。質問文章の詳細・曖昧さは、回答者が回答できるかどうか（自分の知識に合っているか）、回答しようと思うかどうか

(不真面目な文章ではないか)といった回答動機に影響を与える重要な要素の一つである。本稿では、質問を投稿する質問者 i の質問努力 $Effort_{Qi}$ を 0 から 1 の実数値として以下のように表現する。この $Effort_{Qi}$ の値が大きいくほど質問者 i は努力して文章を記述する (質問 $Info_{Qi}$ を生成する) ことを表す。

$$0 \leq Effort_i^Q \leq 1 \quad (式 2)$$

質問 ($Info^Q$)

質問者 i によって投稿される質問 $Info_{Qi}$ を、質問者 i の疑問 $Prop_{Qi}$ をもとに、質問努力に応じて曖昧さが生じるものとして以下のように表現する。

$$Info_i^Q = [x_1, x_2, \dots, x_k] = [x_n \in \{0, 1, *\}] \quad (式 3)$$

ここで、ビット列における*は情報の曖昧さを意味するものとし、 $Prop_{Qi}$ の各ビット列が $1 - Effort_{Qi}$ の確率で*に変化するものとする。具体的には、各ビット列の値は質問者の疑問 $Prop_{Qi}$ と対応し、ある知識単位に対する質問者が求める知識特性の値を意味している。また、あるビット列が*であるということは、質問者が投稿した質問において、*に対応する知識単位に関する記述が欠けており、この箇所に関しては回答者が努力してあらゆる知識特性を網羅的に記述することによってうめられる状態の部分であることを表現している。

返報閾値 (Th^Q)

質問者 i が、回答者 j に対して返報を行うかどうかは、質問者 i の疑問 $Prop_{Qi}$ と、回答者 j の回答 $Info_{Aj}$ とのマッチ度が、質問者 i の返報閾値特性 Th_{Qi} (0 から 1 の実数値) を越えるかどうかによって決定するものとする。この Th_{Qi} の値が大きいくほど、質問者 i は回答とのマッチ度が高い回答者にしか返報を行わないことを表す。

$$0 \leq Th_i^Q \leq 1 \quad (式 4)$$

ここで、マッチ度の計算は、 $Prop_{Qi}$ と $Info_{Aj}$ のハミング距離を 0 から 1 の実数値に

正規化したもの（ハミング距離/ビット幅 k ）とする。

回答者のモデル

回答知識 ($Prop^A$)

質問に対する回答は回答者の知識に基づいて投稿される。そこで、回答者 i の回答知識 $PropAi$ は、質問者の疑問 $PropQi$ と同様に、ある知識単位に対する知識特性の集合として以下のように表現する。

$$prop_i^A = [x_1, x_2, \dots, x_k] = [x_n \in \{0,1\}] \quad (\text{式 5})$$

回答努力 ($Effort^A$)

回答者が質問に回答する際、質問者の曖昧な質問文章に対してどの程度努力して回答文章を記述して回答するかを表現する特性として、回答者 i の回答努力 $EffortAi$ を以下のように表現する。この $EffortAi$ の値が大きいほど、回答者 i は努力して文章を記述することを表す。

$$0 \leq Effort_i^A \leq 1 \quad (\text{式 6})$$

回答閾値 (Th^A)

回答者 i が、質問者 j の質問 $InfoQj$ に対して回答を行うかどうかを、回答者 i の回答知識 $PropAi$ と質問者 j の質問 $InfoQj$ のマッチ度に対する回答閾値 $ThAi$ として以下のように表現する。

$$0 \leq Th_i^A \leq 1 \quad (\text{式 7})$$

ここで、マッチ度の計算は、 $PropAi$ と $InfoQj$ のハミング距離を 0 から 1 の実数値に正規化したもの（ハミング距離/ビット幅 k ）とする。なお、ビット列における*は、情報の曖昧さを表現しているため、距離 1 として計算する。

回答 ($Info^A$)

回答者 i の、質問者 j の投稿した質問 $InfoQj$ に対する回答 $InfoAi$ を以下のように表現

する。

$$Info_i^A = [x_1, x_2, \dots, x_k] = [x_n \in \{0,1\}] \quad (式 8)$$

ここで、回答者 i の回答努力 $Effort_{Ai}$ の値が大きいほど、回答者 i は質問者 j の質問 $Info_{Qi}$ に近づけた文章を記述するものとする。具体的には、回答者 i は質問者 j の質問 $Info_{Qj}$ の k 列のビット列中の*について、 $Effort_{Ai}$ を超えるまでのビット列の割合 (*のビット列/ビット幅 k) に対し、質問者 j の質問 $Prop_{Qj}$ の各ビット列に近づけた値 (値のコピー) の回答 $Info_{Ai}$ を生成するものとする。なお、 $Info_{Ai}$ の各ビット列は、回答者の回答知識 $Prop_{Ai}$ 、質問 $Info_{Qi}$ 、質問者の疑問 $Prop_{Qi}$ と対応しており、ある知識単位に対する回答者の知識特性を表している。また、質問 $Info_{Qi}$ の*の部分に対して、回答者が質問者の疑問 $Prop_{Qi}$ に近づけた値を生成するという行為は、質問において情報が欠けている箇所に対して、回答者が努力してあらゆる知識特性を網羅的に記述することによって質問者の疑問に合った回答を記述するということを表現している。

援助性 (S)

我々は、[三浦 08]での回答者の動機因子を参考に、回答者 i が回答という行動のみで得られる満足 of 利得を S_i (0 から 1 の実数値) として表現する。

$$0 \leq S \leq 1 \quad (式 9)$$

3.3.3. 報酬制度のモデル化

効用は、金銭的なものと非金銭的なものとに分類できる[Fehr 99]。質問者と回答者の知識取引による効用としての報酬を考えたとき、報酬は金銭的な報酬(お金、ポイント)と非金銭的な報酬(感謝・お礼の言葉)とを合わせたものとして表現できる。ここで、金銭的報酬についてはコミュニティの報酬制度として外部から操作可能な変数である。

そこで、本研究では、知識取引における非金銭的報酬を一定とし、Q&A サイトの制度として操作可能な金銭的報酬の大きさを金銭取引性 PT (Pecuniary Transaction) という値として表現する。この値が大きいほど、知識取引における金銭的報酬が大きいコミュニティ制度であることを表す。なお、現実でのサービスにおける報酬を考えた場合に

は、報酬の値の上限に際限はないと考えられる。しかし、本研究における報酬のモデル化においては、質問者が支払う報酬は、最大でも回答を得ることによる利得以上の値は支払わないと仮定し、これに即した抽象的なモデル化とした。具体的には、質問者が払う報酬 PT の値は、回答から得られる利得（質問者の質問と回答とのマッチ度）の範囲である $[0, 1]$ の間としてモデル化する。

$$0 \leq PT \leq 1 \quad (\text{式 } 10)$$

3.3.4. 利用者の効用のモデル化

本節では、3.1～3.3 節であげた知識取引のモデル、利用者のモデル、報酬制度のモデルをふまえ、質問・回答のやり取りといった知識取引から得られる利得とコストをもとに、質問者 i と回答者 j の効用の式を以下のように定式化する。

質問者の効用 (U^Q)

$$U_i^Q = \text{match}(\text{prop}_i^Q, \text{Info}_j^A) - \text{Effort}_i^Q - PT \quad (\text{式 } 11)$$

式(11)における質問者 i の効用 U_i^Q は、回答 j から得られる回答 Info_j^A とのマッチ度の利得関数 $\text{match}()$ から、質問作成コスト Effort_i^Q を引き、さらに回答へのお礼にかかる返報コスト PT を引いた値として表現する。なお、マッチ度の関数 $\text{match}(x, y)$ は、ビット列 x と y のハミング距離を正規化した 0 から 1 の実数値とする（ビット列 x と y のハミング距離/ビット幅 k ）。

回答者の効用 (U^A)

$$U_j^A = PT - \text{Effort}_j^A + S_j \quad (\text{式 } 12)$$

式(12)における回答者 j の効用 U_j^A は、質問者から得られる返報利得 PT から、回答

作成にかかるコスト $Effort_{Ai}$ を引き、回答をすること自体の援助利得を意味する援助性 S_j を足した値として表現する。

3.3.5. コミュニティ環境（質問者と回答者の同質性）

Q&A の利用は、様々なコミュニティ環境での利用が考えられる。例えば、Yahoo!知恵袋のような多様なカテゴリーのもとで多様な知識を持った利用者が参加するコミュニティ、または特定の専門的な話題を扱い同様の知識傾向を持った利用者が集まるコミュニティ（UNIX コミュニティ等）、あるいは、知識傾向が異なる異分野間での相補的なナレッジマネジメントとして Q&A を用いる行うコミュニティなどが考えられる。

本研究では、これらのコミュニティ環境の違いを、質問者と回答者の同質性として定義し、質問者の疑問 $PropQ$ と、回答者の回答知識 $PropA$ の値の発生確率を変化させることでこれを表現する。具体的には、 $PropQ$ の各ビットが 1 を取る確率を表現した確率変数を X 、 $PropA$ の各ビットが 1 を取る確率を表現した確率変数を Y とする。ここで、 X と Y は互いに独立な確率変数とし、 X の期待値を E_X 、 Y の期待値を E_Y とすると、 $PropQ$ と $PropA$ の一致の期待値 z は、以下の式で計算する。

$$z = (1 - E_X)(1 - E_Y) + E_X E_Y \quad (\text{式 13})$$

本研究では、質問者と回答者の同質性が標準で中間的なコミュニティを同質性 $z=0.5$ 、($E_X=0.5$ 、 $E_Y=0.5$)、同質性が高いコミュニティを $z=0.7$ 、($E_X=0.2$ 、 $E_Y=1/6$)、同質性が低いコミュニティを同質性 $z=0.3$ 、($E_X=0.2$ 、 $E_Y=5/6$)としてコミュニティ環境の違いを表現する。

3.3.6. 利用者の離脱・参入

実際の Q&A サイトにおいては、特定のユーザの限定的な利用だけでなく、参加するモチベーションが低くなったユーザの離脱や、新たな利用者の参入といった現象が起こりうる。本稿では、このような現象を表現するために、参加する動機が低くなったユーザ、つまり Q&A サイトにおいて獲得できる効用が低くなったユーザは離脱して、新たなユーザと入れ替わるといった現象としてモデル化する。具体的には、質問者と回答者それぞれにおいて、直近 X 期の効用を累計し、獲得効用の低い下位のいくらかのユー

ザが新規のユーザと入れ替わることでこれを表現する。

3.4. シミュレーション実験

本節では、提案したモデルの妥当性検証と、報酬制度の導入効果、コミュニティ環境（質問者と回答者の同質性）による報酬効果の違いについてシミュレーション実験を行う。

3.4.1. 実験概要

本シミュレーション実験の目的は、報酬制度が、Q&A サイトの成果である回答の数、回答の質、質問のしやすさにおいてどのような効果があるかといったことを明らかにすること、また、異なるコミュニティ環境による報酬の効果の違いを分析することである。

以下にシミュレーションの実験条件を示す。シミュレーションステップについては、每期、以下の Step1～Step3 をシミュレーション期間 t まで繰り返すものとする。

Step1. 行動選択

- 質問者は、質問をコミュニティに投稿する
- 回答者は、コミュニティ上の質問に対して自身から見たマッチ度の計算をもとに、回答閾値を超えた質問に対して回答を行う
- 質問者は、自身の質問のすべての回答に対して返報するかどうかを、自身から見たマッチ度の計算をもとに、返報閾値を超えた回答に対し返報を行う

Step2. 効用計算

- 質問と回答者は、3.3.4 節であげた式(11)、式(12)をもとに、それぞれ効用 U_Q 、 U_A を得る

Step3. 離脱・参入

- 質問者と回答者は、直近 X 期の累計効用を累計し、下位 $Y\%$ の利用者はコミュニティから離脱する
- 離脱で減った利用者だけ、質問者と回答者に新規利用者が参入する

実験で用いたパラメータを（表 1）に示す。なお、シミュレーション実験の結果は、乱数のシードを変化させた 100 回のシミュレーション試行の平均値を結果とした。また、ステップ数に関しては、500 ステップの段階で集団の特性が収束していることから 500

ステップ後の結果について掲載している。

表 1: シミュレーション実験パラメータ (表中の Q は質問者、A は回答者を表す) }

	パラメータ	実験で用いた値
シミュレーション条件	シミュレーション期間 t	500
	エージェント数 n	200 (質問者数:100、回答者数:100)
	直近効用の累計期間 X 期	5
	各期のエージェントの離脱割合 $Y\%$	10%
	知識特性のビット幅 k	20
報酬制度	知識の金銭取引性 PT	[0.1, 0.1, ..., 1.0]
コミュニティ環境	Q と A の知識特性の類似性 z	[0.3, 0.5, 0.7]
エージェント固有特性	Q の努力特性 $Effort^Q$ (0~1.0)	[一様乱数]
	Q の返報閾値 Th^Q (0~1.0))	[一様乱数]
	A の努力特性 $Effort^A$ (0~1.0)	[一様乱数]
	A の回答閾値特性 Th^A (0~1.0))	[一様乱数]
	A の援助性特性 Si (0~1.0)	[一様乱数]

3.4.2. モデルの基本的挙動

提案したモデルの基本的挙動として、モデル上での利用者の特性と、現実の Q&A サイトでの利用者の特性が、3.3.4 節で示した効用の要素と対応しているかどうかを比較した (図 3-1)。

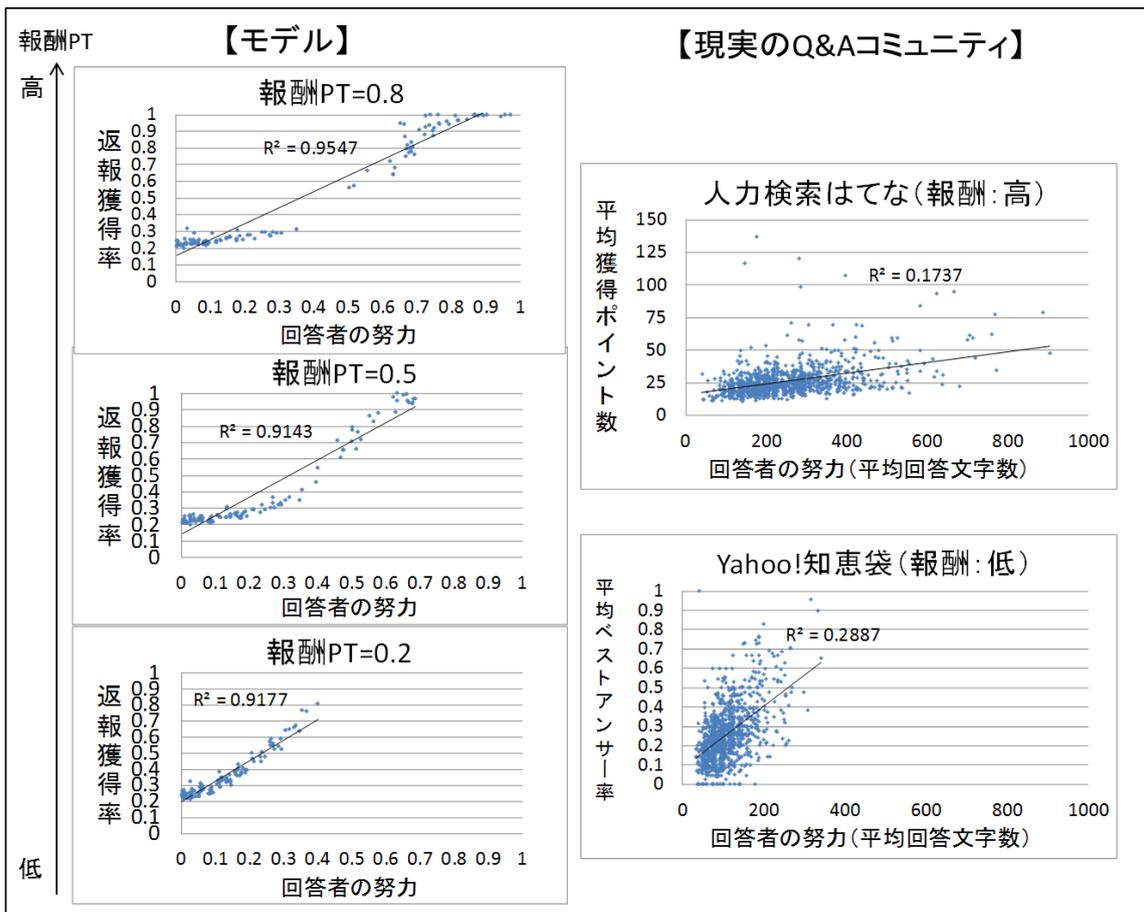


図 3-1 : モデル上の回答者と、現実の Q&A サイトの回答者の特性の比較

(図 3-1) 中のプロットは、一人の回答者の回答努力と返報獲得 (平均) の値である。Yahoo!知恵袋のデータに関しては、国立情報学研究所の提供しているデータを用いており、データの最新日である 2005 年 10 月において回答数 10 回以上の回答者 1,000 人をランダムで抽出したものをプロットしている。人力検索はてなのデータについては、2009 年 11 月において回答数 10 回以上の回答者 1,000 人をランダムで抽出したものをプロットしている。なお、質問者の特性の比較については、現実のデータにおいて質問

者の主観である回答の質（マッチ度）を数値化することが難しいため、回答者の特性について比較を行った。また、人力検索はてなのデータプロットに関して、y 軸の項目については平均獲得ポイント数としている。これは、人力検索はてなでは、ほとんどの回答に対して少なからず質問者のポイント付与があるために、プロット上で返報獲得率が 1 に偏ってしまうことから、返報獲得率よりもよりの確な返報を反映した平均獲得ポイント数を項目として用いた。

（図 3-1）に示す比較の結果、モデルと現実の利用者特性の一致が確認された。3. 4 節の回答者の効用モデルでは、回答努力は効用に対してコスト項であるため、努力をせずに利得を得ようとする戦略が優位であると考えられる。しかし、図 1 の結果では、回答努力の高い回答者の返報獲得が高くなっている。この結果をもたらすメカニズムは、回答者にとっての回答努力は回答作成の労力コストとなるが、一方で回答労力の高い回答は質問者からの返報率を高める効果があることにより、結果としてより高い利得を得ることが可能になっているためである。現実の Q&A サイトにおいても、回答努力を文章量、利得をベストアンサー率として観察してみると、シミュレーション結果と同様の結果を得ている。

この結果より、本研究での知識取引と利用者の効用のモデルが、現実の Q&A サイトの利用者の利用行動を表現したモデル化となっているといえる。

3.4.3. 報酬制度の導入効果

回答の数

(図 3-2) に、報酬 PT を変化させたときの、回答の数 (1 つの質問への平均回答数) の質問者平均を示す。この結果より、報酬 PT を高めると、どのコミュニティ環境においても回答の数は低下することが分かった。

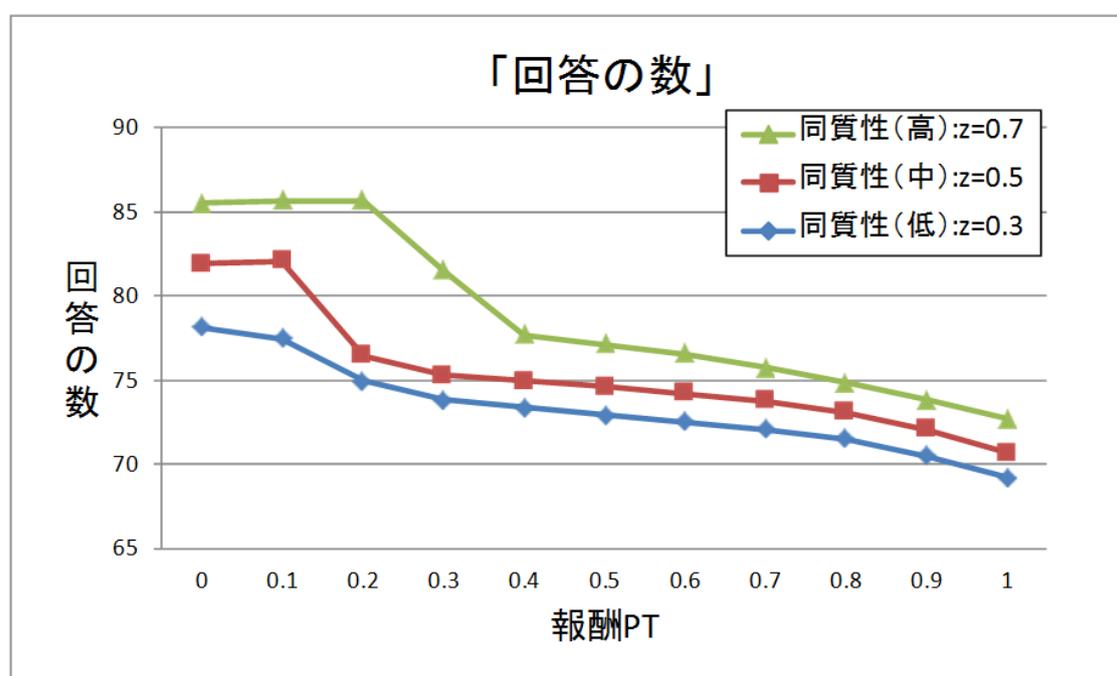


図 3-2 : 報酬 PT を変化させたときの、回答の数の変化

回答の質

(図 3-3) に、報酬 PT を変化させたときの、回答の質（質問者の知識特性と回答との平均マッチ度）の質問者平均を示す。この結果より、報酬 PT を高めると、どのコミュニティ環境においても回答の質は向上することが分かった。

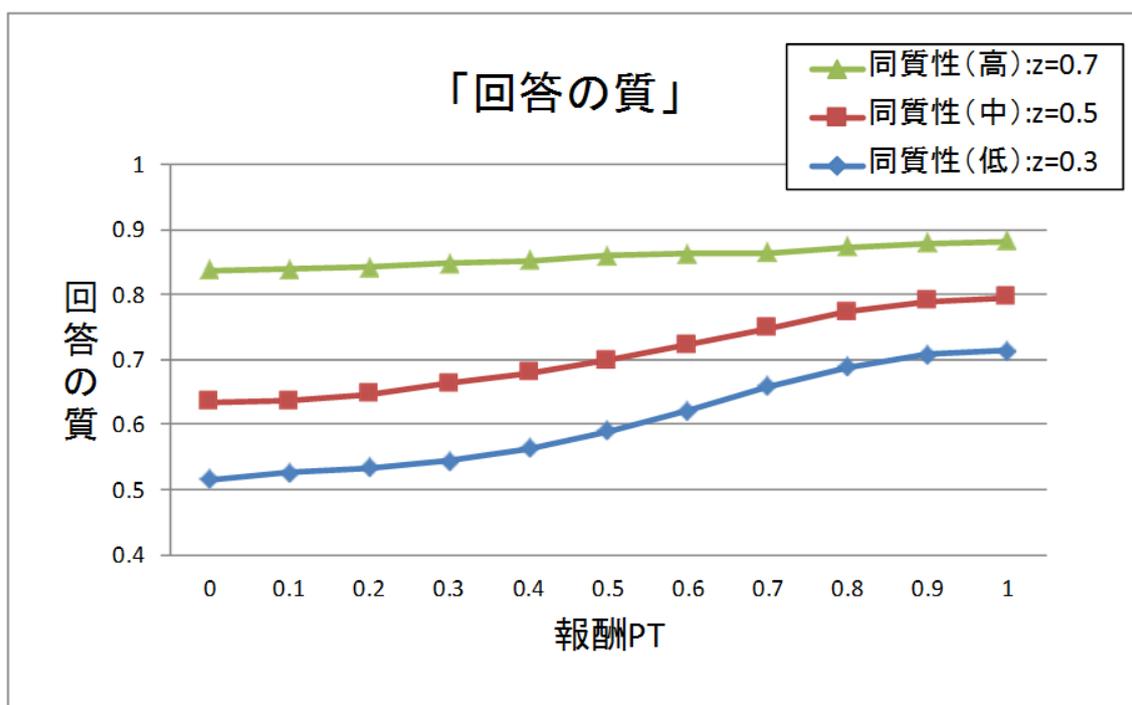


図 3-3： 報酬 PT を変化させたときの、回答の質の変化

質問努力（質問のしやすさ）

（図 3-4）に、報酬 PT を変化させたときの、質問者の質問努力の変化を示す。この結果より、同質性 $z=0.5$ （質問者と回答者の同質性が中間のコミュニティ環境）、同質性 $z=0.7$ （質問者と回答者の同質性が高いコミュニティ環境）では、報酬 PT を高めるにつれて質問努力が低下していることが分かった。ただし、同質性 $z=0.3$ （質問者と回答者の同質性が低いコミュニティ環境）においては、報酬 PT が低い制度の場合において（PT が 0.0~0.2 までの間において）、報酬 PT が中間的なときよりも質問努力が低いことが分かった。

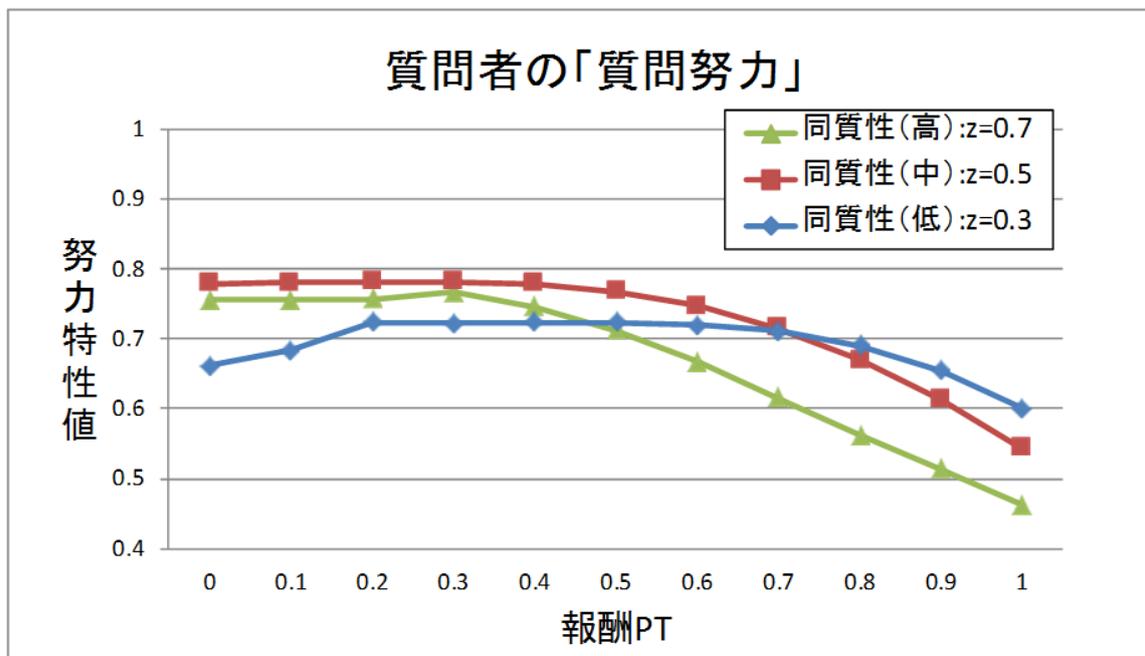


図 3-4： 報酬 PT を変化させたときの、質問努力の変化

また、(図 3-5) に、同質性 $z=0.5$ (質問者と回答者の同質性が中間のコミュニティ環境) における、報酬 PT を変化させたときの質問者の返報閾値と質問努力、回答者の回答努力の関係を示す。同様に、(図 3-6) に同質性 $z=0.3$ (質問者と回答者の同質性が低いコミュニティ環境)、(図 3-7) に同質性 $z=0.7$ (質問者と回答者の同質性が高いコミュニティ環境) での結果を示す。この結果より、報酬 PT の高さに応じて、返報閾値と回答努力が増加していることが分かる。また、質問者の質問努力が、回答者の回答努力の増加に反比例して低下していることが分かる。

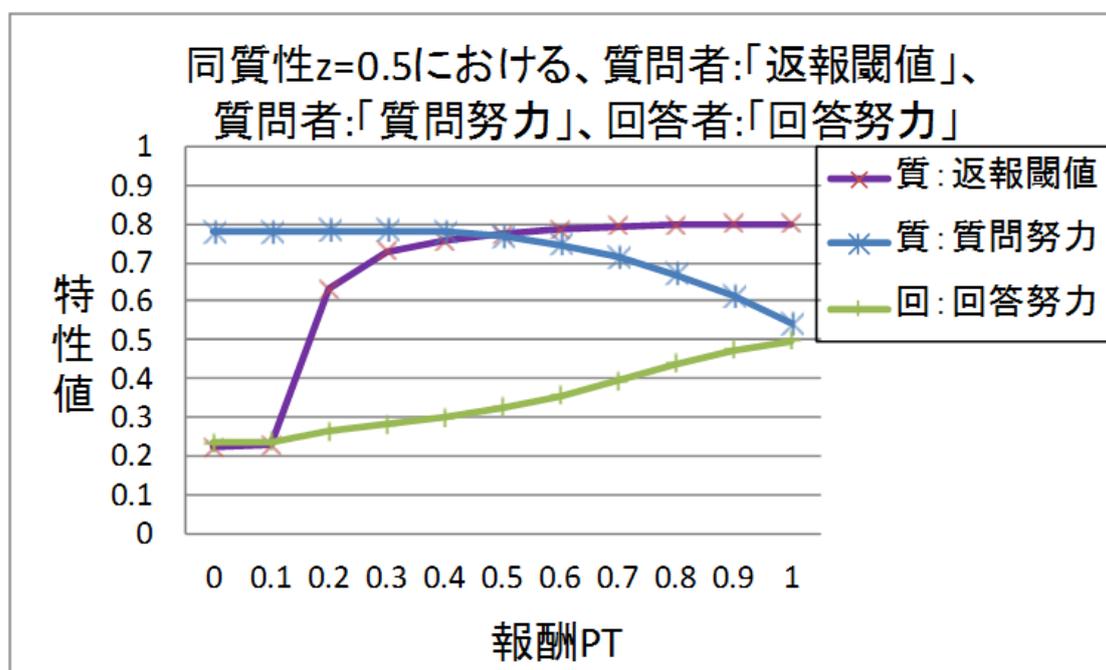


図 3-5: 同質性 $z=0.5$ (質問者と回答者の同質性が中間のコミュニティ環境) における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係

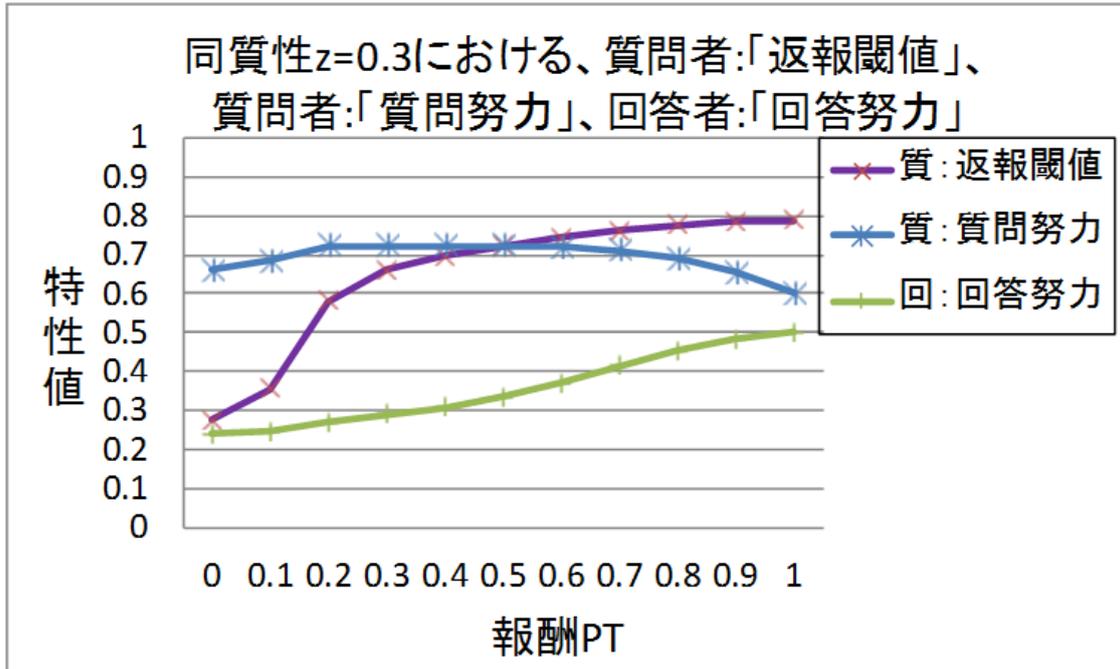


図 3-6: 同質性 $z=0.3$ (質問者と回答者の同質性が中間のコミュニティ環境)における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係

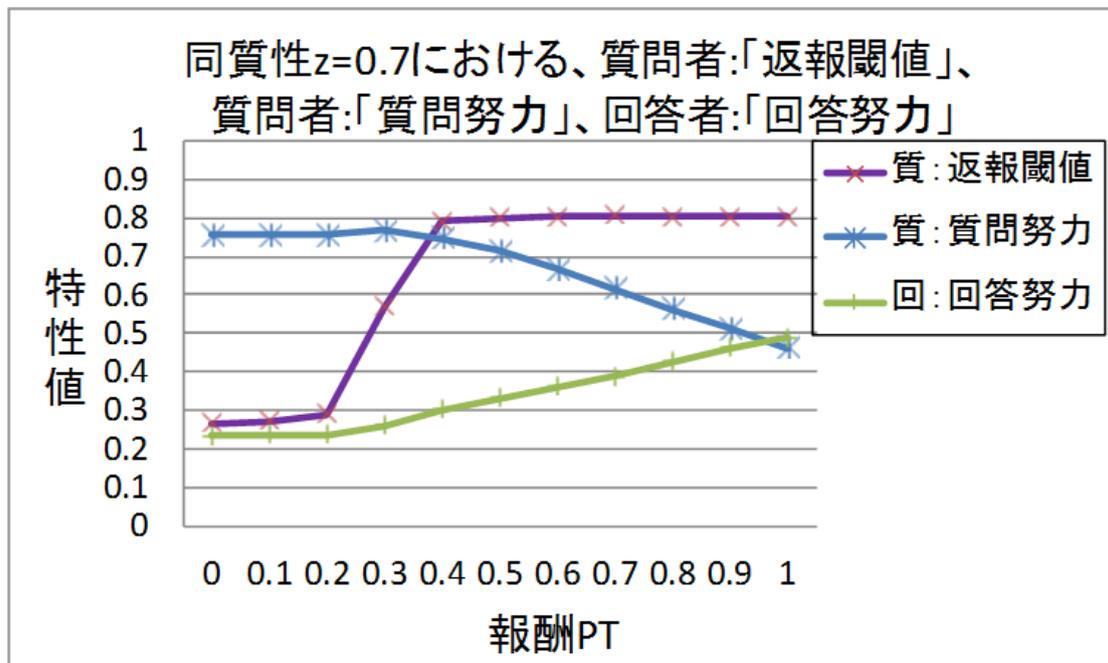


図 3-7: 同質性 $z=0.7$ (質問者と回答者の同質性が中間のコミュニティ環境)における、報酬 PT を変化させたときの返報閾値、質問努力、回答努力の関係

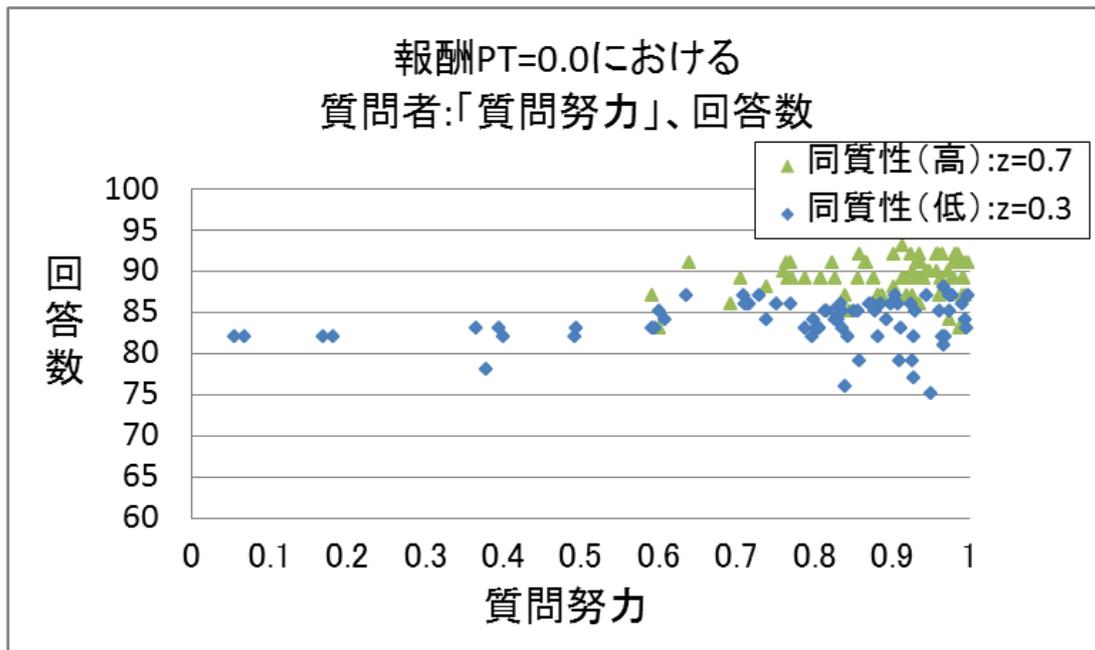


図 3-8： 同質性 $z=0.3$ と同質性 $z=0.7$ のコミュニティ環境での、報酬 $PT=0.0$ における質問努力と回答数の関係の違い

(図 3-8) に、図 4 で同質性 $z=0.3$ のコミュニティが他のコミュニティ (同質性 $z=0.5$ 、 0.7) と異なり、報酬 PT が低い場合において PT が中間的な場合よりも質問努力が低くなっている要因を表した結果を示す。(図 3-8) は、報酬 PT が低い ($PT=0.0$) 制度の場合における、同質性 $z=0.3$ のコミュニティにおいてどのような質問と回答のマッチングが行われているかを検証するために、質問者の質問努力に対応する回答数をプロットした図である。なお、比較のために、同質性 $z=0.7$ の場合のプロットも同時に示している。図 8 の結果より、同質性 $z=0.3$ のコミュニティにおいては、質問努力が高い質問者に対してのみ回答が集まるのではなく、質問努力が低い質問者でも回答が集まっているという違いがあること、また、質問努力の値が同質性 $z=0.7$ のコミュニティに比べて低いことが分かった。

3.5. 考察

3.5.1. 回答の数と質

(図 3-2) と (図 3-3) の結果より、報酬 PT を高めると、回答の数は減少するが、回

答の質は向上することが分かった。コミュニティ環境による効果の違いとしては、全体傾向としては同じであるが、質問者と回答者の同質性 z が大きいコミュニティほど、回答数の絶対値と、回答の質の絶対値が高くなっているといった違いがみられた。この結果については、 z が大きいコミュニティほど元々質問と回答がマッチしやすい環境であるため、正当な結果であるといえる。

(図 3-5) は、報酬 PT を高めることによって回答の数と質に変化が生じるメカニズムを説明した図である。本モデルにおいて、報酬 PT は質問者にとってコストであるために、報酬 PT が大きい場合においては、返報閾値の低い質問者よりも高い質問者（低いマッチ度の回答に対して返報する質問者よりも、高いマッチ度でのみ回答する質問者）が大きな利得を得ることとなり、このような質問者がコミュニティに存続することとなる。また、これにより、回答者においても、質問に対してよりマッチした回答を行う回答者（回答努力の高い回答者）の方が質問者の返報を獲得することができ、このような回答者がコミュニティに存続することとなる。これらのメカニズムは、(図 3-5) における質問者の返報閾値、回答者の回答努力より確認することができる。

3.5.2. 質問努力（質問のしやすさ）

(図 3-4) の結果より、全体傾向として、どのコミュニティ環境においても、報酬 PT を高めると PT が低い制度のときよりも質問努力が低いコミュニティが形成されることが分かった。ここで、質問努力とは質問者がどれだけコストをかけて質問作成を行うかといったことを表す。現実的には、質問者の質問努力の値が低いほど質問時に手をぬいても回答が得られるコミュニティ（「何でも質問して OK」）、逆に値が高いほど手をぬいた質問では回答が得られないコミュニティ（「きちんと調べてから質問してくれ」）であることを意味する。

(図 3-5) は、このような質問努力の低下が生じている様子を示す結果である。(図 3-5) の結果より、質問者の質問努力が、回答者の回答努力の増加に反比例して低下していることが分かる。これは、3.5.1 節で述べたメカニズムによって回答者の努力が高くなれば、質問者にとっては、あまりコストをかけて正確な質問をしなくてもマッチした回答が集まるため、質問努力をかけない回答者が得をするといったメカニズムが働いた結果と理解できる。この結果については、実際の Q&A サイトの人力検索はてなでのポイント制の期待する効果（「過去の Q&A をみれば分かるのに」と言われない効果[折田 09]）を支持する結果である。

(図 3-8) の結果より、同質性 $z=0.3$ のコミュニティにおいては、質問努力が高い質

問者に対してのみ回答が集まるのではなく、質問努力が低い質問者でも回答が集まっているという違いがあること、また、質問努力の値が同質性 $z=0.7$ のコミュニティに比べて低いことが分かった。その原因を推測すると、同質性 $z=0.7$ のコミュニティでは、知識傾向が類似する回答者との接触の確率が高いため、高い質問努力でコストを払うことになっても、回答者が回答しやすい質問が投稿される状況になることで。このため、回答数がより集まることの利得獲得の効果が大きくなり、質問努力の高い質問者がコミュニティに存続したと考えられる。一方、同質性 $z=0.3$ のコミュニティでは、知識傾向が類似する回答者とのコミュニティ内での接触確率が低いため、高い質問努力でコストを払って回答者が回答しやすい質問を投稿するよりも、質問努力が低くコストを払わないままで、たまたま知識傾向の類似した回答者とのマッチを得ることによる利得獲得の影響が大きい。このため、質問努力が低い質問者でもコミュニティに存続できる環境であったと考えられる。特に、報酬 PT が低い場合には、マッチングによる利得効果が大きいためにこの影響が強く出ていると考えられる。以上のことから、同質性 $z=0.3$ のコミュニティでは、同質性 $z=0.7$ のコミュニティよりも、質問者の質問努力が低くなっていると考える。

3.5.3. 報酬制度の導入方針への知見

3.5.1 節と 3.5.2 節の結果より、Q&A サイトにおける報酬制度の導入方針を検討する。

(図 3-2)、(図 3-3)、(図 3-4) の結果より、同質性 $z=0.5$ (質問者と回答者の同質性が中間的なコミュニティ環境) では、回答の数を重視するのであれば、報酬 PT は低く、回答の質と質問のしやすさを重視するのであれば報酬 PT は高く設定するといった導入方針が有効であるといえる。

一方、同質性 $z=0.7$ (質問者と回答者の同質性が高いコミュニティ環境) では、同質性 $z=0.5$ と同様な導入方針であるが、回答の質については、報酬 PT を高めることによる向上はそれほど見込めないため、質問のしやすさとしての効果を目的とした導入方針が有効であるといえる。

また、同質性 $z=0.3$ (質問者と回答者の同質性が低いコミュニティ環境) では、報酬 PT を高めることによって、回答の質の向上と、質問のしやすさにおいては効果が現れるが、質問のしやすさを重視するのであれば、報酬 PT が低い制度の場合でも質問がしやすい環境であるため、無理に報酬 PT を高めて回答の数を低下させるよりも、報酬 PT が低い方が効率的な報酬設定であるといったことがいえるだろう。

3.5.4. 本シミュレーションの限界と制約

なお、シミュレーション実験では、質問者の割合が 50%（全エージェント数 200 の内、質問者のエージェント数 100）の場合の結果のみを掲載したが、これ以外の割合の場合（質問者の割合が 70%、または 30%）においても回答数の絶対値に変化はあるものの、報酬 PT による全体の結果（回答の質、回答の数、質問努力）の傾向に変化はないことを確認している。また、離脱者の割合に関して、実験結果では全体の 10%の場合の結果のみを示したが、離脱者の割合が大きい場合では、高い利得を得ているエージェントでも離脱しやすくなるために、報酬 PT 変更による効果（回答の質、回答の数）が弱まることを確認している。具体的には、離脱者の割合が全体の 50%以上の場合には、報酬 PT が高い場合でも回答の質の向上はほぼ見込めない結果となる。このことから、離脱・参入の激しいコミュニティよりも、Q&A の利用において利益を得ている参加者が継続的に利用しているコミュニティほど、報酬の導入による効果は大きくなると推測できる。

ただし、本シミュレーションの制約として、サイトにおける参加者数の規模が一定であること、また、離脱・参入の質問者・回答者の比率が一定としていることがあるため、本研究で得られた報酬制度の導入方針に関して、現実への一般化には限界がある。今回のシミュレーションモデルにおいては、報酬制度が利用者行動にもたらすメカニズムを理解することを容易にするために、サイトの参加者規模が一定であり、かつ質問者・回答者のそれぞれが一定の割合で離脱・参入するモデルでの分析を行った。しかし、より現実への適用を重視した報酬制度の導入方針を得るためには、現在のモデルをベースを拡張し、サイトにおける参加者数の規模の増減や、質問者・回答者の比率が動的に変化するといった多様な環境下での分析に関しても検討する必要がある。

3.6. 結論と今後の課題

本稿では、Q&A サイトにおける報酬制度の違いが、コミュニティにどのような効果をもたらすのか、また、コミュニティ環境によって効果にどのような違いがあるのかといった報酬制度の導入方針の知見を得ることを目的に、Q&A サイトのエージェントベースモデルを提案し、シミュレーション実験を行った。その結果、そのコミュニティが何を重視するのか（回答の数か、回答の質か、質問のしやすさか）によって、有効な報酬制度は異なっているという知見を明らかにした。具体的には、回答の数を重視するの

であれば報酬 PT は低く、回答の質と質問のしやすさを重視するのであれば報酬 PT は高くすることが有効であることを明らかにした。

今後の課題としては、報酬制度のモデル化において、非金銭的報酬による回答者のモチベーション向上についても検討する必要があるだろう。提案したモデルでは、報酬はコミュニティが設定できる金銭的報酬についてのみ扱ったが、現実の Q&A サイトにおいて、長期間参加する回答者にとっては、金銭報酬よりも名声などの他者評価も重要な回答モチベーションの要素として考えられる。このような観点でのモデル化も、今後の Q&A サイトの有用な制度設計を検討していくうえでの重要な課題であると考えられる。

4章. Novelty 向上のための推薦手法の設計

4.1. 研究背景

近年、インターネットの急速な普及により、膨大な情報を誰でも容易に取得できるようになった。しかしその反面、利用者が本当に望む情報を取得することが非常に難しくなっており、有用な情報をいかに選定するかという情報フィルタリング技術の重要性が高まってきている。そのなかで推薦システムと呼ばれる技術は、多くの情報の中から利用者の望む情報をフィルタリングする有用な技術として、Amazon.comをはじめ多くの商用サイトなどで広く用いられ大きな成功をおさめている [Schafer 01]。

従来の推薦システム研究において、推薦の有用性は、推薦がいかにユーザの好みに合っているかという推薦の正確さとして捉えられ、その評価指標である予測精度（精度・再現率、MAE） [Herlocker 04] の向上が重視されてきた。しかし、最近ではその高い推薦精度が災いして、好みではあるが似たようなアイテムばかり推薦されてしまい、結果的にユーザはそれだけでは満足しないという結果も報告されている [Ziegler 05、McNee 06]。例えば、推薦システムを備えた EC サイトにおいて、一旦ジブリ映画作品の DVD を購入したとき、その後しばらくは他のジブリ作品の DVD が積極的に推薦される。これは推薦の正確さという観点では適切な推薦であるが、推薦された作品を既にユーザが見ていた場合、これらの推薦はユーザにとって退屈な推薦であり、長く続けば推薦への飽きに繋がり、結果的にはユーザはそのサイトを利用しなくなるということが考えられる。

今後の推薦システムの課題として、推薦システムの有用性を、従来の正確さの観点だけでなく、ユーザに新たな価値の発見を提供するといった正確さ以外の観点から新たに検討することの重要性が指摘されている [Smyth 01、 Herlocker 04、 McNee 06]。例えば、ユーザにとって目新しさのある推薦 [Herlocker 04] を行うことは、既に知っているアイテムが推薦されることによる推薦に対する飽きの問題を解消し、ユーザの新たな嗜好の開拓に繋がると考えられる。

本研究では、推薦の有用性を従来の正確さの観点から捉えるのではなく、推薦の目新しさという観点に注目し、目新しいアイテムを推薦することでユーザ満足度を向上させる推薦手法の構築を目的とする。本研究において推薦の目新しさ (Novelty) とは、ユーザが「今まで知らなかったが、推薦されて (アイテムのタイトル・概要文・画像を見

て) 興味を持ったアイテム」を推薦できることと定義する。我々は、嗜好傾向が類似するアイテムを複数推薦するよりも、嗜好傾向の異なるアイテムを複数推薦する方が、ユーザの知らないアイテムを推薦する可能性が広がるという仮定にもとづき、推薦リストの嗜好傾向の多様化手法により Novelty を向上させる手法を新たに提案する。

本章の構成を以下に示す。4.4.2 節では先行研究と本研究の位置づけとして、従来の推薦システムのアルゴリズム、推薦の評価指標、推薦の正確さ以外の向上を目的とした推薦システム研究について述べ、本研究の位置づけと目的を示す。4.4.3 節では、本研究で提案するトピックの多様化を用いた推薦手法について、嗜好傾向に基づいたトピック抽出の方法、トピックを用いた推薦の多様化手法について詳細を述べる。4.4 節では、提案手法の有効性を評価するための評価実験について述べる。4.4.5 節において実験結果に対する考察、4.4.6 節においてまとめ、4.4.7 節で今後の課題を述べる。

4.2. 先行研究と本研究の位置づけ

4.2.1. 従来の推薦システムのアルゴリズムと評価

推薦システムの実現方法は、大きく分けて、コンテンツベースフィルタリング方式と、協調フィルタリング方式とに大別される[Adomavicius 05, 神島 08]。

コンテンツベースフィルタリング方式は、アイテムの属性情報から抽出した特徴情報と、ユーザの嗜好を表現した特徴情報を比較し、それらが類似するアイテムを推薦対象アイテムとして選定する手法である [Pazzani 96, Mooney 99]。この方式は、推薦の質が利用するユーザの数やアイテムの数に影響されず、運用の初期段階でも比較的安定した推薦が行えるといった利点がある。しかし、アイテムによっては特徴を解析することが困難である場合もあり、特徴情報をどのように抽出し表現するかといった問題がある。また、アイテムの特徴のみを用いて推薦対象を決めるため、推薦アイテムが似かよったものになりやすく、意外な推薦が行われ難いといったことが問題点としてあげられる。

協調フィルタリング方式は、ユーザの嗜好を評価履歴という形で記録し、推薦対象ユーザと似た評価履歴をもっているユーザの評価履歴をもとに、ユーザの嗜好を推測し、推薦対象となるアイテムを選定する手法である [Resnick 94]。この方式では、ユーザの評価にもとづいてユーザ間、あるいはアイテム間の類似性を求めるため、アイテムの内容の特徴解析を行わない。そのため、どんな種類のアイテムに対しても同じ手法が適用可能であるといった利点がある。類似度計算において、ユーザ間の類似度を計算する場合はユーザベース協調フィルタリング [Resnick 94]、アイテム間の類似度を計算する場

合はアイテムベース協調フィルタリングと呼ばれている [Sarwar 01]。それぞれの特徴として、ユーザベースでは、ユーザ間の類似度を用いるため、コンテンツベースに比べ Serendipity の高い推薦が行われる余地が大きいという利点がある。一方、アイテムベースではユーザベースよりも予測精度が高く計算量が少なくて済むという利点があるが、一時的な個人化までしか行えず、似たアイテムばかり推薦されてしまうということが実験的に示されている [McNee 06]。また、協調フィルタリング全体の欠点として、各ユーザの嗜好を把握するためには、多くのアイテム評価情報が必要であるといった問題がある [Mooney 99]。そのため、これら互いの方式の欠点を補うために、コンテンツベースフィルタリング方式と協調フィルタリング方式の統合したハイブリッド方式も提案されている [Balabanovic 97、Claypool 99]。この他にも、Web ページの推薦などにおいて、ユーザの嗜好データの不足を Web ページに付与されるタグといった抽象データによって補うことで解決しようといった Folksonomy を用いた新たな推薦方式も提案されてきている (17)。しかし、これらの研究の多くは、推薦の正確さを高めることを目的とした研究であり、正確さ意外の観点については重視されていない。

推薦システムによる推薦結果を評価する方法に関して、[Herlocker 04] や [神島 07] は評価方法の分類や具体的な評価指標について述べている。推薦の評価方法は、オフラインで行う場合と、オンラインで行う方法とに分けられる。オフラインでの評価とは、ユーザから事前に集めた嗜好のサンプルデータと、そのサンプルデータより生成した予測データを比較し、その一致程度を評価する方法である。一般的には交差検証法による評価方法が多く用いられている。この方法は、事前にユーザの嗜好のサンプルデータさえあれば評価が行えるため、実際の調査を行う必要がなく、調査のためのコストが少ないという利点を持つ。ただし、交差検証法による評価は、サンプルデータと予測するデータは同じ分布から得られることを仮定しているため、厳密な予測精度の評価ではないという欠点がある。一方、オンラインでの評価とは、ユーザに実際に推薦システムを利用してもらい、推薦が適切かどうかを評価する方法である。この方法は、ユーザからの直接的な評価であり、実際の運用に近い評価であるという利点があるが、調査のためのコストが大きいという欠点がある。

推薦の具体的な評価指標は、推薦の正確さの指標とそれ以外の指標に大別できる。

推薦の正確さは、ユーザの関心のあるアイテムを推薦できるかという点で最低限備えておくべき推薦の規準である [Swearingen 01]。従来の推薦の評価では、多くの研究がこの正確さの向上を重視している。正確さの評価指標としては、精度・再現率のような評価指標が用いられる。オンライン評価の精度は、推薦リスト中の好みのアイテム (5

段階評価の場合、上位 2 段階の評価のアイテムなど)の割合として評価する方法である。この指標は、個々の推薦アイテムへの評価を用いた評価指標であるが、推薦リスト全体へのユーザ満足度とも高い相関があり [Swearingen 01]、最も基本的な評価指標とされている。オフライン評価の精度と再現率は、情報検索システムの評価指標としてよく知られており、協調フィルタリングによる推薦リストの正確さの指標として用いられている。この他、個々の推薦アイテムの予測の正確さの評価として、予測評価値と実際のユーザの評価値がどれだけ近いかを評価する平均絶対誤差 (MAE) などがある。

しかし、現在では推薦の有用性を正確さの指標のみで測るだけでは不十分であることが指摘されている [Herlocker 04, McNee 06]。推薦システムにおいて、ユーザの好みのアイテムを推薦するという正確さの観点は重要であるが、好みであってもユーザが既に経験したアイテム (見た、あるいは読んだことがあるアイテム) だけを推薦しても、ユーザにとって本当に有用な推薦とはいえない。このため、正確さ以外の指標について、異なる観点で様々な指標の提案がなされている。代表的な評価指標としては、Novelty や発見性、Serendipity などがある。推薦の Novelty とは、ユーザが関心をもち、かつそれがまだユーザの知らないものである推薦の目新しさのことを指す [Herlocker 04]。また、発見性 [清水 08] は、推薦リスト中の知らないアイテムの割合を評価する指標である。また、Serendipity は、目新しさに、思いがけなさ、予見のできなさ、または意外さの要素が加わった概念として示されている [神嶋 07]。ただ、Serendipity はユーザの感情的な面に影響される部分が大きいため、定量的に示すことが難しい評価とされている [McNee 06]。いくつかの提案指標として、推薦リスト内アイテムの類似性 (Intra-List-Similarity) [Ziegler 05] や、評価対象である推薦システムの予測結果とプリミティブな推薦システムの予測結果との差異を意外性として評価した意外性 [村上 07] などの指標が考えられている。

4.2.2. 推薦の正確さ以外の向上を目的とした研究

推薦の正確さ以外の向上を目的とした研究としては、清水ら [清水 08]、Ziegler ら [Ziegler 05]、村上ら [村上 09] の研究がある。

清水ら [清水 08] は、ユーザの「知らないアイテム」を推薦することを推薦の発見性と捉え、ユーザのアイテムへの既知・不既知の情報を用いて協調フィルタリングを行うことにより発見性を向上させる推薦手法を提案した。この結果、既存の協調フィルタリングによる推薦よりも、ユーザの知らない・好みのアイテムを多く推薦できることを

オフラインの Novelty の評価実験により示した。しかし、実際の運用の場面を考えたとき、知らないアイテムに対してユーザから「知らない」という情報を明示的に得ることは難しいと考えられるため、ユーザ労力の観点においては課題が残る。よりユーザ労力の少ない方法で有効な効果が出せる手法について考える必要がある。また、推薦の評価がオフラインで行われる Novelty 精度による評価のみであるため、オンラインでのユーザの直接的な評価によっても推薦手法の有効性を示す必要がある。

Ziegler ら[Ziegler 05] は、アイテムに付与しているジャンルや著者などの静的なカテゴリ情報 (Amazon.com におけるカテゴリ分類情報) を利用して、推薦リスト内のアイテム類似度を計算する指標として Intra-List-Similarity を提案し、この類似度が低くなるように推薦アイテムを選定する多様化の手法を提案した。この結果、ある程度の多様化を行った推薦リストによって最も高いユーザ満足度が得られたことをオンラインでのユーザ満足度の実験で示した。しかし、この手法の問題点として、カテゴリという既に定義されている静的なカテゴリ情報があることを前提としているため、適用範囲が本や DVD などのカテゴリ付けが容易なアイテムに限定されてしまう点がある。また他の問題点として、アイテム間には、静的なカテゴリだけでは捉えきれない類似関係が存在すると考えられる。作品・テーマの類似性という観点でいえば、例えば、同じジブリ作品の映画というカテゴリであっても、多くの人が好む「となりのトトロ」と、賛否両論分かれる「ゲド戦記」は必ずしも強い類似関係があるとは限らない。また、話題の類似性という観点では、ヒットした映画などにおける認知度の違いなどが考えられ、これらはジャンルという静的なカテゴリだけでは捉えきれない。話題性は動的に変化し、テーマ・思想などを一人で定義することは分類に恣意性を生じさせる。これらのことから、多様化に用いるカテゴリについて、テーマ・話題の類似性などを考慮したカテゴリを人手により設計するには、大きな労力と困難さを有する。

また、村上ら [村上 09] は、嗜好モデルの推薦リストと習慣モデルの推薦リストをマージすることで推薦の意外性を向上させる手法を提案し、TV コンテンツを対象に実験を行い、その有用性を示した。しかし、個人の習慣といったものは、利用するコンテンツの特性に影響を受けていると考えられる。そのため、習慣モデルの作成には、コンテンツの特性を考慮したうえでの適切なモデルが必要になるため、そのモデル作成のための分析は困難であり、適用範囲も限られてくると考えられる。

4.3. 動的なトピック分類に基づく推薦リストのトピック多様化推薦

本研究では、ユーザのアイテム評価という行動履歴から動的に変化する嗜好傾向に基づき、共に評価されやすいアイテム群のカテゴリ（これをトピックと呼ぶ）を自動的に作成し、それに基づく多様化によって Novelty の高い推薦、すなわちユーザが「知らなかったが、推薦されて興味を持ったアイテム」を推薦する手法を提案する。Novelty の高い推薦を行うことは、ユーザの推薦に対する飽きの問題を解消し、ユーザの新たな嗜好の開拓に繋がると考えられるため、推薦の正確さ以外の観点として重要な要素といえる。

提案する手法において、トピックの多様化によりユーザの知らなかったアイテムを推薦するという考えは、特定の嗜好傾向をもったユーザに対して同じトピックのアイテムを推薦しても、既にユーザはそのアイテムを知っている可能性が高いという仮定に基づいている。例えば、推薦リスト内のアイテムが特定の嗜好傾向で偏っていたとき（ジブリ作品映画のみの推薦リストの場合など）、ユーザがその推薦アイテムのどれかに興味があり既に見ていた、あるいは知っていた場合、ユーザは推薦リスト内の同じ嗜好傾向を持つその他のアイテムについても知っている可能性が高いと予想できる。そのため、特定のトピックのアイテムを複数推薦するよりも、異なるトピックのアイテムを複数推薦する方が、ユーザの知らないアイテムを推薦する可能性が高いと考えられる。

以上の考えのもと、我々は、嗜好傾向の異なる複数のアイテムを推薦リストとしてユーザに提示することにより、ユーザの知らないアイテムを推薦する。また、推薦リストの作成段階において、トピックの選定とトピック内からのアイテムの選定に予測評価値を用いた優先順位を設けることで、推薦の正確さを確保する。これにより、Novelty の高い推薦を行う。

なお、本研究で目標とする推薦は、ユーザの新たな発見や興味の開拓に繋がるための推薦である。このためのアプローチとして、普段のユーザの嗜好とは全く関係ないコンテンツの推薦を行うことで「知らないコンテンツ」を推薦する方法が考えられるが、このような推薦は意外であっても、必ずしもユーザが受容する（好みと評価する）とは限らないと考えられる。このため、ユーザが実際に興味を持ったか（好みと評価したか）といった観点が重要になり、そのためには、ユーザの興味範囲のある程度の絞り込みが必要である。以上をふまえ、提案手法では好みの守備範囲を、既存の予測精度の高い協調フィルタリングによってある程度絞り込み、そこから、多様なトピック（全ユーザの興味傾向にもとづくクラスタ）が多様になるコンテンツを推薦することで、既存研究において議論された Novelty 評価指標（推薦リスト中の、知らない・好みのコンテンツの

割合)を向上させるというアプローチをとる。

4.3.1. アルゴリズムの概要

提案する推薦システムの全体図を(図 4-1)に示す。提案手法は以下の3つのステップに分かれる。

- step1. 嗜好傾向に基づくアイテムのトピック分類
- step2. ユーザごとにパーソナライズされた推薦リストの作成
- step3. トピックを用いた推薦リストの多様化

step 1 では、ユーザの嗜好傾向を反映したトピックを作成する。トピックの作成には、ユーザの嗜好行動を表した全ユーザのアイテム評価履歴を用いる。具体的には、全ユーザのアイテム評価行列から「このアイテムを評価しているユーザは、このアイテムも評価している」といったアイテム間の評価の共評価関係をネットワークとして表現し、これをクラスタリングすることにより、嗜好傾向が近いアイテムをトピックとして分類する。

step 2 では、多様化を行う前段階の処理として、あらかじめユーザごとにパーソナライズされた推薦候補となるアイテムを選定しておく。本研究では、既存のユーザベースの協調フィルタリングを用いて、パーソナライズされた推薦リストを作成する。このリストをベースとし、step 3 の多様化の処理においてトピックが多様になるようにアイテムを選定する。

step 3 では、step 1 で得られたアイテムのトピック情報を用いて、step 2 で得られた推薦リストのトピック多様化の処理を行う。具体的には、トピックが複数に及び、かつ step 2 の協調フィルタリングで得られる予測評価値が高くなるようなアイテムを選定することで、多様化を行った新たな推薦リストを作成する。

上記の方法によって、ユーザが「知らなかったが、推薦されて興味を持ったアイテム」の推薦を、ユーザのアイテム評価情報という少ない情報のみで実現する。以下の節では、それぞれの step における処理の詳細を述べる。

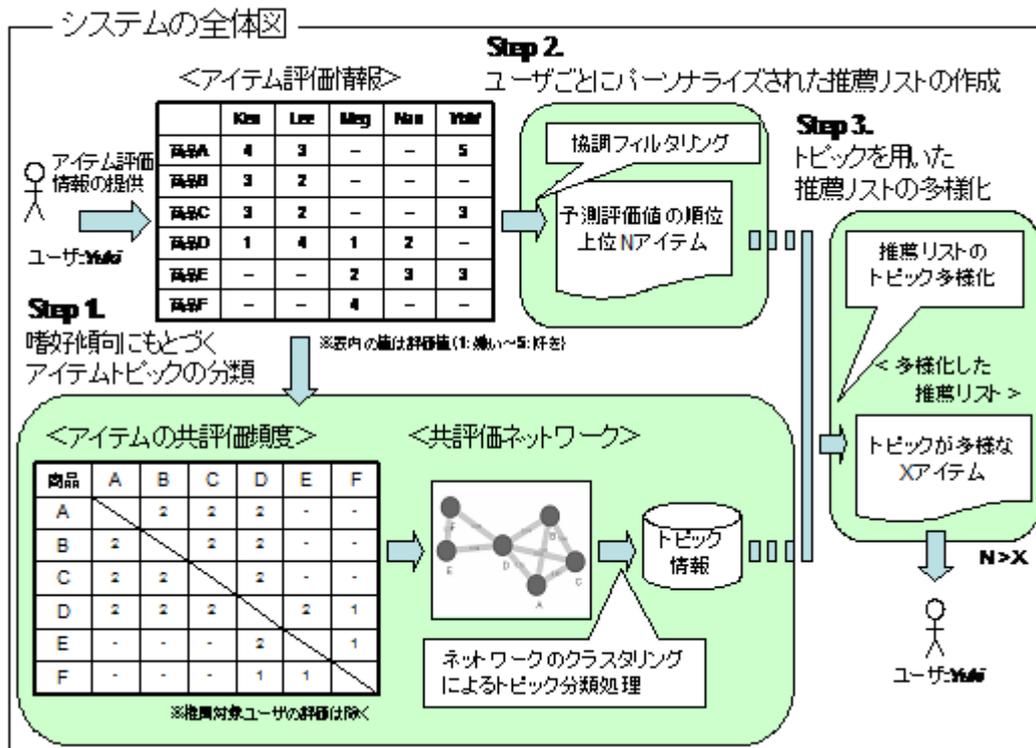


図 4-1：提案する推薦システムの全体図

4.3.2. 嗜好傾向に基づくアイテムのトピック分類（Step 1）

我々は、ユーザーのアイテム評価履歴という行動履歴からアイテムの共評価関係をネットワーク化し、このネットワークをクラスタリングすることにより嗜好傾向を反映したトピックの作成を行う。

ネットワークのクラスタリングをもとにしたカテゴリズに関する研究として、グラフ分析を用いた文書集合からのトピック抽出[戸田 07]や、論文の共著関係からの研究者コミュニティの抽出[松尾 05]、また、SNS のリンク関係を利用したコミュニティの分析[湯田 06]など、現在多くの研究が行われている。これらは、潜在的なコミュニティやトピックの発見、また、時事的に変化する情報を組織的に表現できる点で有効な手法である。このことより、我々は、アイテムの共評価関係をネットワークとして表現し、これをクラスタリングする手法が、嗜好傾向にもとづいたトピックを抽出する方法として有効な手法であると考えます。

以下、トピック分類に用いる共評価ネットワークの作成方法、そのネットワークのクラスタリングによるトピック分類方法についての詳細を述べる。

共評価ネットワーク

トピック分類に用いるネットワークには、アイテムをノード、エッジにアイテム間の共評価の頻度の重みを設けた重み付き無向ネットワーク（以下、共評価ネットワーク）を用いる。

ネットワークのクラスタリングによるトピック分類

トピックの分類は、共評価ネットワークを Newman らのアルゴリズム [Newman 04a] でクラスタリングすることにより行う。

Newman らは、クラスタリングの性能を評価するモジュール性 Q という指標を提案し [Newman 04b]、この指標を用いたクラスタリング手法を提案した。ここで Q 値が示すものは、「モジュール内でのノード間にリンクが存在する確率の実測値－ランダムネットワークと仮定した場合のモジュール内におけるリンクの割合の理論値」であり、モジュール内のリンクが密で、かつモジュール間のリンクが疎である場合にこの値は大きくなる。上記の手法は、重みなし無向ネットワークにおけるクラスタリング手法であるが、重み付きネットワークへの応用も可能である [Newman 04c]。また、エッジに重みを設けることでクラスタのサイズが平均化され、クラスタリングの結果が改善されるという結果も示されている [安藤 06]。

Newman のアルゴリズムはエッジが密なほど同じクラスタに所属することになる。更に、エッジに共評価の頻度を重みとして与えることで、エッジが密で高い共評価頻度が高いものがより同じクラスタに所属しやすくなる。共評価頻度が高いということは、それらのアイテムはあるユーザから見ると共に知られている可能性が高いということであるので、同じクラスタに所属しているアイテムは嗜好傾向が近いアイテム群となる。つまり、トピックとは、そこに所属するアイテムを互いに知っている評価者が多いことになるので、トピックが異なれば、知らない可能性が高くなる。

以上より、我々は、エッジにアイテムの共評価の頻度の重みを設けた重み付きネットワークに対し、Newman らのクラスタリング手法を用いることで、嗜好傾向が類似するアイテムのグループをトピックとして分類することができると思う。

4.3.3. ユーザごとにパーソナライズされた推薦リストの作成 (Step 2)

提案手法では、多様化を行う前にあらかじめユーザごとにパーソナライズされた推薦候補アイテムを選定しておく。このリストをベースとし、後の多様化の処理において、このリストからトピックが多様になるようにアイテムを選定する。

本研究では、既存のユーザベース協調フィルタリングを用いて、ユーザごとにパーソナライズされた推薦リストを作成する。

ユーザベースの協調フィルタリングによる推薦は、ユーザと好みの似たユーザグループが好きなアイテムをそのユーザに推薦するという手法である。手順としては、類似度計算と予測評価値計算の2つの手順で行われる。類似度計算では、推薦を受けるユーザと似た嗜好を持っているユーザを推定するために、ユーザ間の類似度を計算する。類似度の計算にはコサイン類似度やピアソン相関係数 [Resnick 94] が用いられる。予測評価値計算では、類似するユーザの評価をもとに、未評価のアイテムに対して、そのアイテムの予測評価値 P を計算する。この値の上位 N アイテムが、Top- N 推薦リストとしてユーザに提示される。以下に、ユーザ U_x の未評価アイテム I_a への予測評価値の計算式(3)を示す。

$$P_{x,a} = ave(r_x) + \frac{\sum_{k \in K} sim(x, k)(r_{x,a} - ave(r_k))}{\sum_{k \in K} |sim(x, k)|} \quad (式 14)$$

ここで、 $ave(r_x)$ はユーザ U_x の投票したすべての評価値の平均値、 $\sum k \in K$ はユーザ U_x と類似度の高い近傍ユーザとする。

上記のユーザベース協調フィルタリングを用いることにより、ユーザごとにパーソナライズされた推薦リストを作成する。

4.3.4. トピックを用いた推薦リストの多様化 (Step 3)

我々は、ユーザにとって Novelty のある推薦を実現するために、複数の異なるトピックから推薦アイテムを選定するトピック多様化アルゴリズムを提案する。多様化の手法としては、4.3.2 で得られたアイテムのトピック情報を用い、4.3.3 で得られた推薦リストの中からトピックが多様になるようにアイテムを選定する。(図 4-2) にトピック多様化のアルゴリズムを示す。

```

procedure TDA( $L_i, X$ ) {
   $L_i^{TD} \leftarrow \{\}$ ;
   $L_i^{CF} \leftarrow L_i$ ;
  if ( $X > |L_i^{CF}|$ ) return  $L_i^{TD}$ ; //  $X$  が範囲外の場合

  while (1)
     $TP \leftarrow \{\}$ ;
     $list \leftarrow \{\}$ ;

    for all  $item \in L_i^{CF}$ 
       $topic \leftarrow \text{getTopic}(item)$ ;
      if (!( $\text{has}(TP, topic)$ ))
         $\text{add}(list, item)$ ;
         $\text{add}(TP, topic)$ ;
      end if
    end for

    for all  $item \in list$ 
       $\text{add}(L_i^{TD}, item)$ ;
       $\text{delete}(L_i^{CF}, item)$ ;
      if ( $|L_i^{TD}| == X$ ) go to end while;
    end for
  end while
  return  $L_i^{TD}$ ;
}

```

図 4-2 : トピック多様化のアルゴリズム

ここで、既存の協調フィルタリングから生成されるユーザ i の N アイテムの推薦リスト L_i から、トピックを多様化し、かつ既存のリスト L_i^{CF} の順位を維持した X アイテムの新たな推薦リスト L_i^{TD} を作成することを考える（ただし、 $X < N$ とする）。多様化アルゴリズムの概要は以下の通りである。

まず、最初の for 内において、既存の推薦リスト L_i^{CF} から各トピック最低 1 アイテムずつ、かつ推薦順位が高いアイテムの順に選んだリスト $list$ を作成する。具体的には、既存の推薦リスト L_i^{CF} の上位から順に 1 アイテム選定し、そのアイテムのトピックを 3。2 で得られたトピック情報を用いて調べ ($getTopic(item)$)、そのアイテムのトピックがまだ選定したことのないトピックであれば ($!has(TP, topic)$)、 $list$ に追加する ($add(list, item)$)。以上の処理により、 $list$ に追加したアイテムは多数のトピックにまたがることになるので、ユーザのまだ知らない可能性の高いアイテムが選定されやすくなる。また、あるトピックから推薦するアイテムを抽出する際に、予測評価値の高い順に抽出することで、ユーザの興味の高いアイテムが選ばれやすくなる。これらによって、Novelty の高い推薦リストを作成することができる。

次に、2 番目の for 内において、作成した $list$ 内のアイテムを多様化リスト L_i^{TD} に加え ($add(L_i^{TD}, item)$)、追加したアイテムは既存のリストから削除する ($delete(L_i^{CF})$)。以上の処理を、 L_i^{TD} のアイテム数が X に達するまで繰り返すことで、推薦リストに複数の異なるトピックを含ませ、かつ予測評価値の高いトピックを優先したリスト L_i^{TD} を作成する。

4.4. 評価実験

提案手法が、Novelty の高い推薦を実現できるかを評価するために、従来の協調フィルタリング手法（以後、CF）との比較実験を行った。

4.4.1. 実験概要

実験対象

実験は、大学生 20 人に対して行った。

データセット

本実験では、アイテム集合として Amazon.co.jp の DVD アイテム 1,000 個、評価値情報として学生 20 人と Amazon.co.jp のレビュー者の情報を用いた。データは Amazon

API¹⁰によって 2007/4/18～2007/4/20 の間にかけて収集し、DVD アイテム 1,000 個は、Amazon.co.jp の全 DVD アイテム数に対する各ジャンルのアイテム数の割合を考慮し、各ジャンルでの売上上位のアイテムを選定した。内訳は、「外国映画：224 個」、「日本映画：49 個」、「アニメ：179 個」、「ミュージック：151 個」、「TV ドラマ：87 個」、「ドキュメンタリー：17 個」、「お笑い：12 個」、「ステージ：26 個」、「ホビー：63 個」、「スポーツ：65 個」、「キッズ：55 個」、「BOX セット：48 個」、「アイドル：24 個」である。上記の 1,000 アイテムに対する Amazon.co.jp のレビュー数は、9,102 レビューであり、この内レビュー数が 1 つしかないユーザを除き、最終的にユーザ数 1,609 人、レビュー数 5,692 のデータを得た。

実験手順

実験の手順を以下に示す。

(1) アイテム評価情報の収集：

ユーザ (学生 20 人) にランダム 30 個の DVD アイテムを提示し、既知・不既知 (1：見た、2：見てないけど知っている、3：知らない) と、評価値 (1：全く見たくない、2：あまり見たくない、3：どちらでもない、4：やや見たい、5：とても見たい) を入力してもらいアイテム評価のデータを得る。また、評価情報の標本として Amazon のレビュー者の情報を利用する。

(2) トピックの分類：

Amazon.co.jp のレビュー者の情報を用いて、アイテムの共評価ネットワークを作成し、これをクラスタリングしてアイテムをトピックに分類する。

(3) 多様化推薦リストの作成：

(1)のデータの内、ユーザの「知っているアイテム (見た or 見てないけど知っているアイテム)」を用いてユーザベースの協調フィルタリングで推薦候補上位 Top-50 アイテムを選定する。そして、(2)のデータと推薦候補の 50 アイテムを用いて、多様化を行った上位 X アイテムの推薦リスト ($X \in [10, 20, 30, 40, 50]$) を作成する。

¹⁰ AmazonWebService : <http://aws.amazon.com/jp/>

(4) オンラインでのユーザ満足度評価 :

(3)で作成した推薦リストをユーザに提示し、各アイテムに対して、既知・不既知 (1: 見た、2: 見てないけど知っている、3: 知らない) と、評価値 (1: 全く見たくない or もう全く見たくない、2: あまり見たくない or もう見たくない、3: どちらでもない、4: ややみたい or たまに見たい、5: とても見たい or また見たい) を入力してもらいユーザ満足度のデータを得る。

(5) オンライン評価分析 :

(4)で得たアイテム評価をもとに推薦リストの評価を行う。本論文では、推薦リスト中の「好みのアイテム」の割合を Precision と定義する。さらに、それを知っていたかどうかという観点から、推薦リスト中の「知っている・好みのアイテム」の割合を Accuracy、推薦リスト中の「知らなかったが、推薦されて興味を持ったアイテム」の割合を Novelty と定義する。

なお、ユーザが知らないアイテムに対しても評価を行えるよう、(1)と(4)でユーザに提示するアイテムは、Amazon.co.jp で用いられている「タイトル」、「表紙画像」、「商品の概要文」を共に提示した。

4.4.2. 実験結果

トピック分類の結果

収集したアイテムの評価情報より、(図 4-3) のアイテムの共評価ネットワークを作成した。このネットワークをクラスタリングした結果、トピック数 355、1 トピックの平均アイテム数 2.82、クラスタリング性能も $Q=0.29$ とモジュール性を持ったクラスタリング結果を得た。トピックに関しては、アイテムのジャンル、シリーズ、あるいは視聴対象者等で似かよっており、嗜好傾向を反映したトピック分類結果が得られている (表 2)。

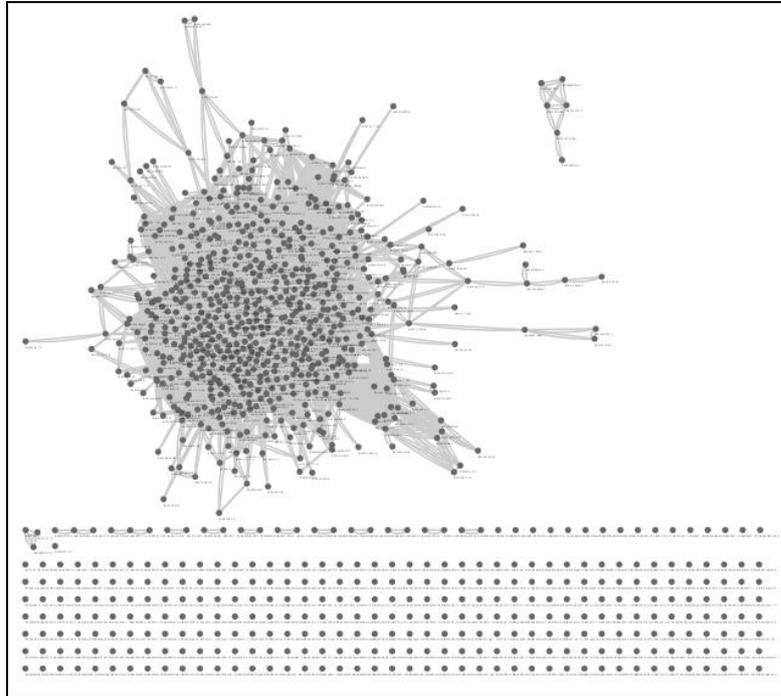


図 4-3 : アイテムの共評価ネットワーク (ノード数 : 1,000 エッジ数 : 9,171)

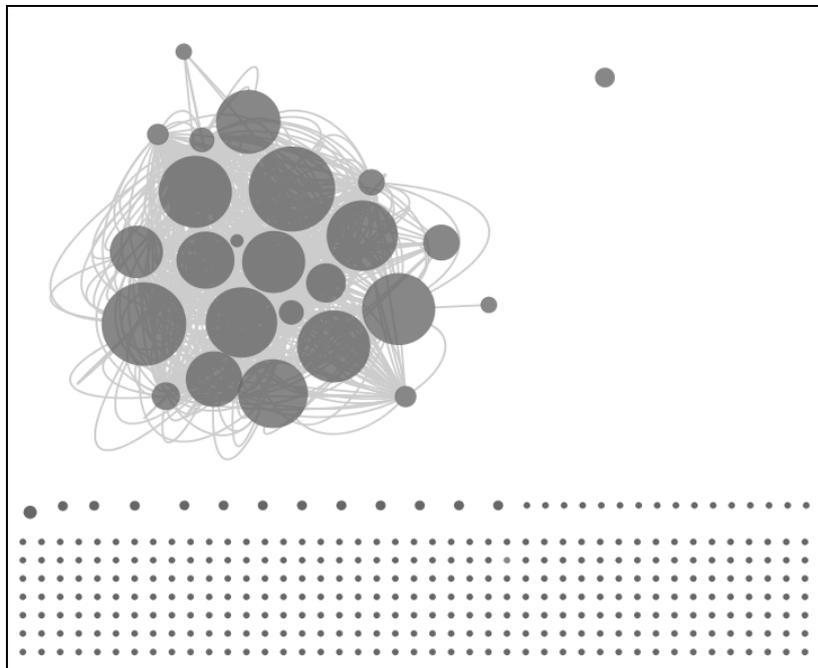


図 4-4 : クラスタリング結果 (ノードの大きさはクラスタ内ノード数の多さを表す)

表 2： 2007 年次のトピック分類結果（一部）

タイトル	
トピック A	All About 東方神起
	KinKi Kids Dome Tour 2004 - 2005 -Font De Anniversary.- (通常版)
	東方神起 HISTORY of JAPAN vol.1
トピック B	DVD版 ダンス・スタイル・ロッカーズ
	ダンス・スタイル・ヒップホップ
	ダンス・スタイル・ファースト
	ダンス・スタイル・プレイカーズ
	ダンス・スタイル・ベイシック
	ダンス・スタイル・ポップン
トピック C	F1 レジェンド スーパーパフォーマンス '87' 95
	FASTER
	NHK うたって おどろんぼ! ~うたとダンスのゆかいななかまたち~
	NHKおかあさんといっしょ最新ソングブック ふしぎはすてき
	TVアニメシリーズ 『ぼのぼの』 DVD-BOX vol.1
	TVアニメシリーズ 『ぼのぼの』 DVD-BOX vol.2
	WE ARE THE CHAMPIONS FINAL LIVE IN JAPAN
	X-ファイル シーズン・フォー DVDコレクターズ・ボックス
	アメリ
	あらしのよるに スタンダード・エディション
	ウィ・アー・ザ・ワールド・20thアニヴァーサリー
	ゴーストライダー1&2
	ゴッドファーザーDVDコレクション
	サウンド・オブ・ミュージック <ファミリー・バージョン>
	さよなら銀河鉄道999 -アンドロメダ終着駅- (劇場版)
	ティファニーで朝食を
	トイ・ストーリー
	トイ・ストーリー2
	となりのトトロ
	ドラゴンクエスト~勇者アベル伝説~ コンプリートDVD-BOX(限定生産)
	ドン・キホーテ
	ファンタジア
	ブラザー・サン シスター・ムーン
	プリティ・ウーマン 特別版
	マトリックス リローデッド
	マトリックス 特別版
	メリーポピンズ スペシャル・エディション
	もののけ姫
	モンスターズ・インク
	ルパン三世 -カリオストロの城
	ロード・オブ・ザ・リング - スペシャル・エクステンデッド・エディション
	銀河鉄道999 (劇場版)
	銀河鉄道の夜
	劇場版NARUTO -ナルト- 大興奮! みかづき島のアニマル騒動(パニック)だってばよ
	紅の豚
	耳をすませば
	超時空要塞マクロス ~愛・おぼえていますか~
	天空の城ラピュタ
	百色眼鏡
	風の谷のナウシカ
	魔女の宅急便

また、提案したトピックがユーザの嗜好に基づいて動的に変化し、それらが既存のカテゴリ情報では捉えられないアイテムの類似性を表現できることを、2007/4/20 時点でのレビューデータを用いて作成したトピックと、1年後のデータを加えたデータ（2008/4/20 時点でのレビューデータ）から作成したトピックとをもとに比較し確認している（表3）。

表3：2007年次と2008年次のトピックの変化（一部）

2007年次のトピック		2008年次のトピック	
ト ピ ク X	硫黄島からの手紙 期間限定版	ト ピ ク Y	武士の一分
	硫黄島からの手紙 (特製BOX付 初回限定版)		大奥 スタンダード・エディション
	武士の一分		ブレイブハート
	父親たちの星条旗 期間限定版		フラガールスタンダード・エディション
	父親たちの星条旗 (特別版)		UDON スタンダード・エディション
	大奥 スタンダード・エディション		009-1 ゼロゼロナインワン vol.6
	岸辺のふたり		009-1 ゼロゼロナインワン vol.5
	華麗なる一族		ト ピ ク Z
	ワンス・アポン・ア・タイム・イン・アメリカ	硫黄島からの手紙 (特製BOX付 初回限定版)	
	ワールド・トレード・センター	父親たちの星条旗 期間限定版	
	ロッキー DTSコレクターズBOX	父親たちの星条旗 (特別版)	
	ユーリ・ノルシュテイン作品集	モンティ・パイソン・アンド・ホーリー・グレイル	
	モンティ・パイソン・アンド・ホーリー・グレイル	デイ・アフター・トゥモロー	
	パプリカ (Blu-ray Disc)	サウンド・オブ・ミュージック	
	トゥルーマン・ショー (通常版)	あしたのジョー-COMplete DVD-BOX	
	ティファニーで朝食を	DVD~Debu Vs Debu	
	ディパーテッド (期間限定版)		
	サウンド・オブ・ミュージック		
	クライング・ゲーム DTSスペシャル・エディション		
	キングダム・オブ・ヘブン ディレクターズ・カット		
エリザベスタウン スペシャル・コレクターズ・エディション			
エラゴン 遺志を継ぐ者 (Blu-ray Disc)			
イントゥ・ザ・ブルー			
イルマール			

(表3) のトピック Y では、「武士の一分」と「フラガール」が同トピックとして分類されている。これらのアイテムは、ジャンルというカテゴリでみた場合、邦画という大きなジャンルでは同じ分類であるが、その内容は「時代劇」と「青春・学園」と全く異なるものであり、現在の Amazon のカテゴリ分類情報のみからでは類似性を確認できない。しかし、これらのアイテムは 2007 年度の日本アカデミー賞の受賞という共通の類似性が存在しており、実際に 2007 年度以降の共評価関係が増加することで、2008 年次では同トピックとして分類できている。また、2007 年次のトピック X 中の「硫黄

島からの手紙」「父親たちからの星条旗」は、2007年次ではアイテム数が多いトピックであったが、2008年次以降に共評価が増加することで、2008年次ではより嗜好傾向に近い独立したトピックとして分類されている。以上のことから、共評価関係からトピック分類を行う提案手法は、既存のカテゴリー分類では表現できない嗜好傾向に基づいた動的に変化するトピックを生成できているといえる。

提案手法のような自動的に分類するというアプローチとは別に、人手で新たなカテゴリーを追加することで対応するという方法も考えられる。しかし、刻々と変化する情報からこれらの類似性を人手で推測するには大きな労力を有することを考えると、自動で分類できることが望ましい方法と考える。

4.4.3. 多様化推薦リストの結果

トピック分類の結果を元に、提案手法によりトピックを多様化した推薦リストを作成した。(表4)に、ある被験者のCFのリストの一部(上位20アイテム)と、それをトピックで多様化させた提案手法のリストの一部(上位7アイテム)の比較を示す。この結果、CFのリストでは推薦上位のアイテムが同トピックのアイテムでかたまりやすい傾向が見られた。一方、提案手法では、推薦上位から各トピックにまたがって推薦アイテムを選定しているため、推薦リストの上位でもトピックが偏ることなく複数の異なるトピックのアイテムを推薦できている。具体的には、CFのリストではジブリ作品の多いトピックCのアイテムが推薦リストの上位にきているが、提案手法のリストではジブリ作品以外のアイテムも推薦できている。

表 4： CF でのリスト（左）と提案手法でのリスト（右）の比較

タイトル	トピック	タイトル	トピック
天空の城ラピュタ	C	天空の城ラピュタ	C
紅の豚	C	ランボー/怒りの脱出	D
ランボー/怒りの脱出	D	悪魔のいけにえ	E
モンスターズ・インク	C	ナイロビの蜂	F
もののけ姫	C	ユージュアル・サスペクツ	G
Turn 8 ラグナセカの青い空	D	ファイナルファンタジーVII アドベントチルドレン	H
魔女の宅急便	C	カーズ	I
悪魔のいけにえ	E
タイガー&ドラゴン「三枚起請の回」	E		
砂と霧の家 特別版12	E		
ルパン三世 - カリオストロの城	C		
耳をすませば	C		
FASTER	C		
ナイロビの蜂	F		
ユージュアル・サスペクツ	G		
ブラックホーク・ダウン	E		
ロスト・イン・トランスレーション	E		
ファイナルファンタジーVII アドベントチルドレン	H		
バタフライ・エフェクト	E		
カーズ	I		
...	...		

また、(図 4-5) に、推薦リスト内のトピックの網羅数を全被験者の平均で表したものを示す。この結果より、図 2 で示したトピック多様化アルゴリズムによって、実際に推薦リスト内のトピックが多様化していることがわかる。具体的には、CF では推薦リストの上位 50 アイテム内における全トピックを網羅するのに 50 アイテムの推薦が必要になるが、提案手法では上位 20 アイテムの段階で全トピックを網羅できていることがわかる。

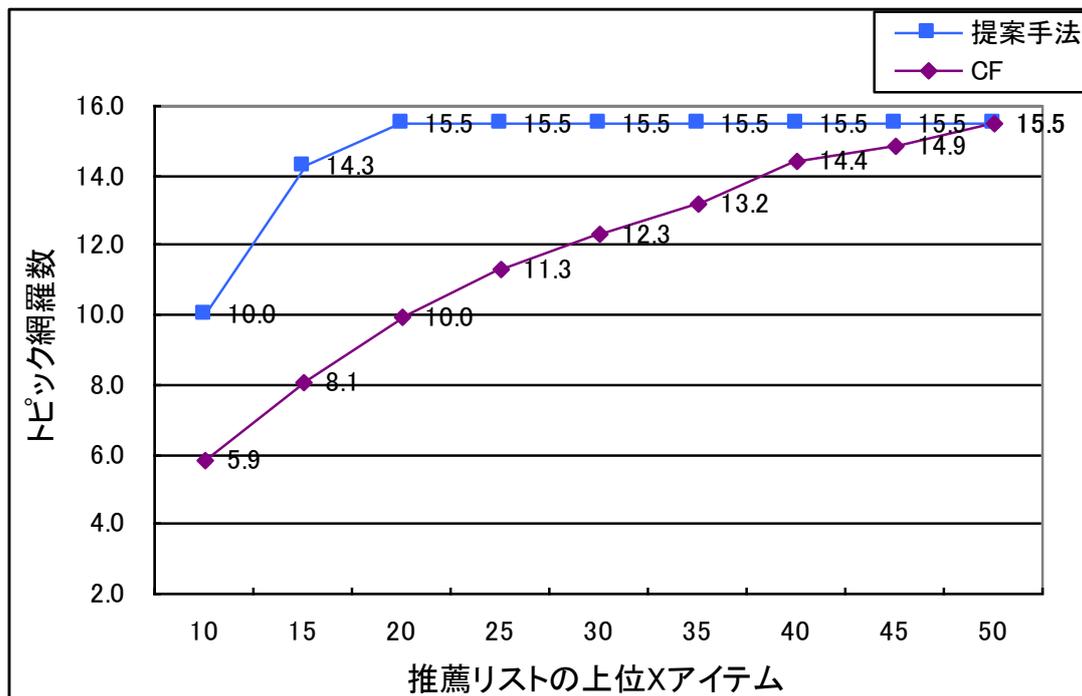


図 4-5： 推薦リストの上位 X アイテム中におけるトピック網羅数

4.4.4. ユーザ満足度評価の結果

(図 4-6) に、推薦リスト上位 X アイテム中のユーザの「知らなかったが、推薦されて興味を持ったアイテム (5 段階評価中の上位 2 段階評価のアイテム)」の割合 (Novelty) を示す。対応のある平均値の差の検定の結果、推薦リストの上位 10 位において 1% の有意差、上位 20 位において 5% の有意差が確認されている。この結果より、提案手法が CF よりも Novelty の高い推薦が実現できていることがわかる。

図 7 に、推薦リストの上位 X アイテム中のユーザの「知っている・好みのアイテム」の割合 (Accuracy) を示す。対応のある平均値の差の検定の結果、有意な差はみられなかった。この結果より、提案手法が CF と同程度のユーザの「知っている・好みのアイテム」を推薦できていることがわかる。

(図 4-8) に、推薦リストの上位 X アイテム中のユーザの「好みのアイテム」の割合 (Precision) を示す。対応のある平均値の差の検定の結果、有意な差は見られなかった。この結果より、提案手法が CF と同程度のユーザの「好みのアイテム」を推薦できていることがわかる。

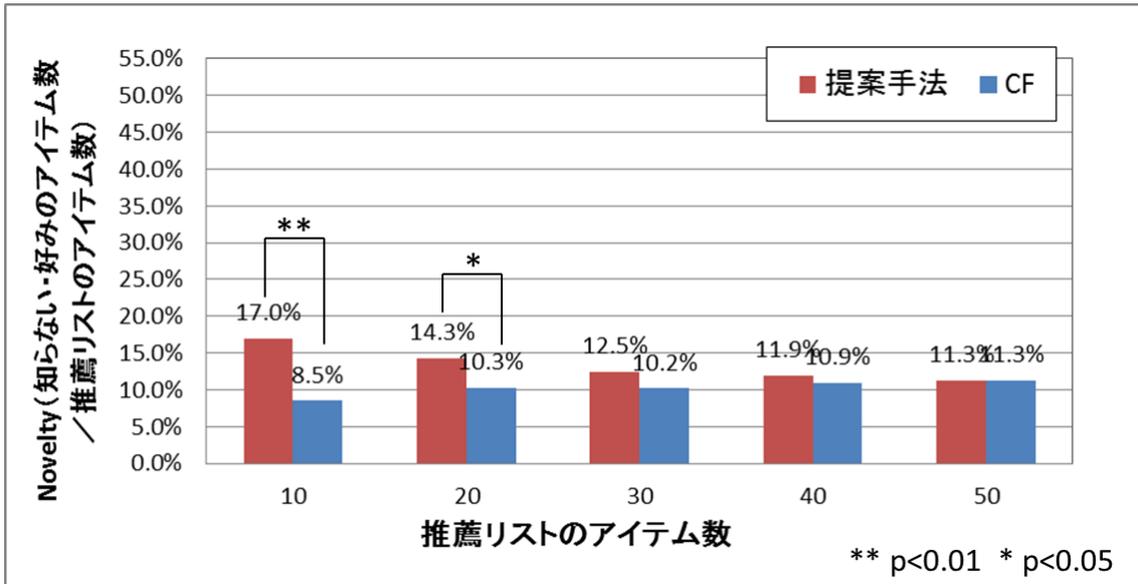


図 4-6 : Novelty : 推薦リストの上位 X アイテム中における「知らなかったが、推薦されて興味を持ったアイテム」の割合

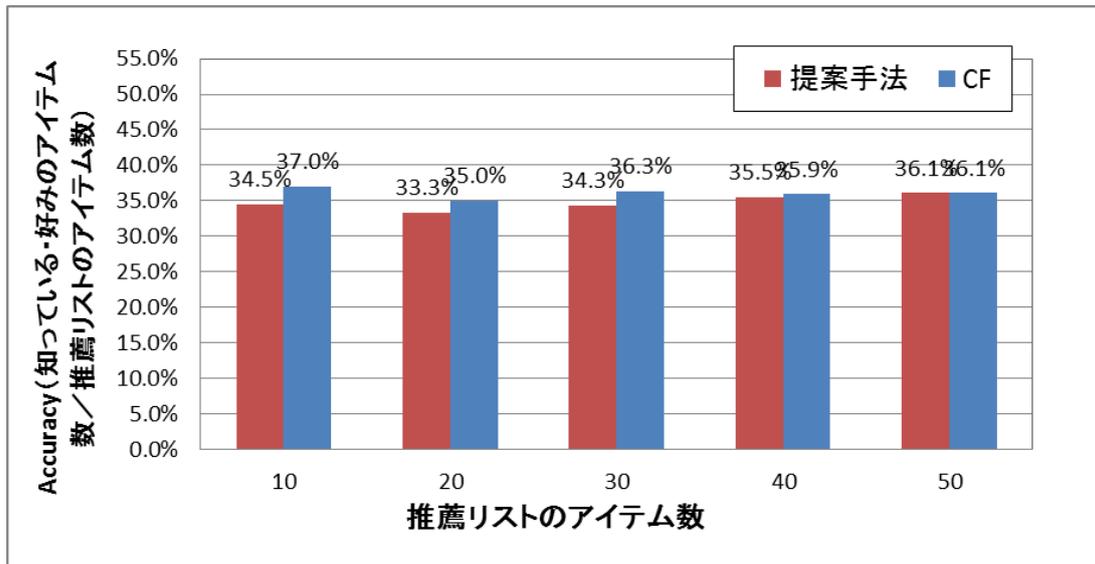


図 4-7 : Accuracy : 推薦リストの上位 X アイテム中における「知っている・好みのアイテム」の割合

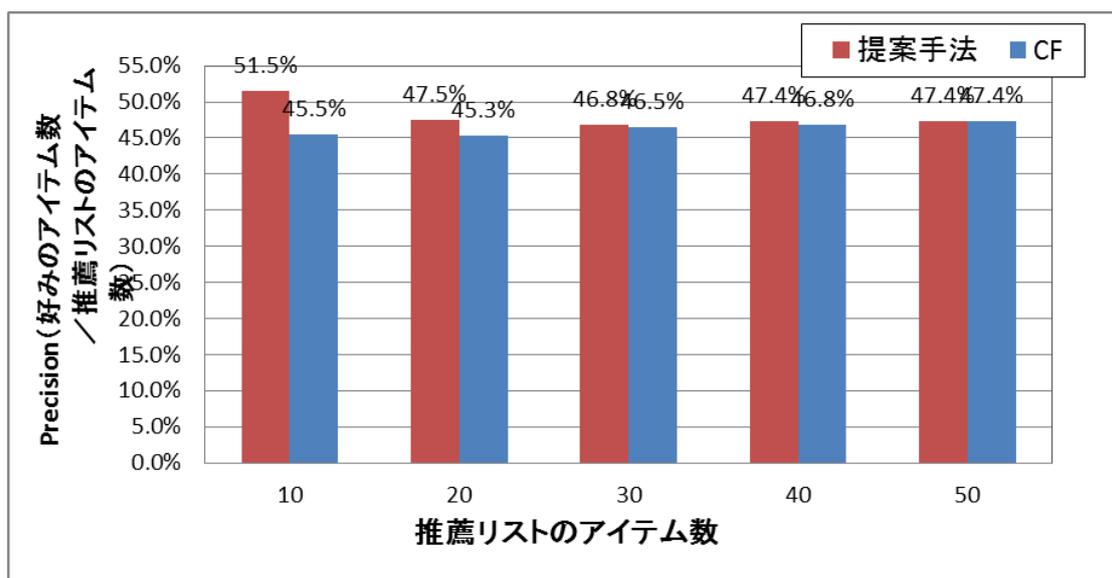


図 4-8 Precision : 推薦リストの上位 X アイテム中における「好みのアイテム」の割合

4.5. 考察

本研究で目指す推薦は、Novelty の高い推薦、つまりユーザが「知らなかったが、推薦されて興味を持ったアイテム」を推薦することである。提案手法においては、上位 50 位までの推薦アイテムの並べ替えを行っているため、アイテム数が多くなるほど差はなくなることは当然である。ユーザは推薦リストの上位より閲覧するため、推薦リストの上位において有意な差が出るのが、効果を測定する上で重要となる。(図 4-6) の対応のある平均値の差の検定結果において、推薦リストの上位 10 位で 1% の有意差、上位 20 位で 5% の有意差が確認されている。このことから、トピックの多様化と予測評価値を用いて優先度の設定する提案手法が、Novelty の高い推薦をするうえで有効な手法であるといえる。

一方、ユーザの「知っている・好みのアイテム」を推薦することに関しては、対応のある平均値の差の検定結果より、提案手法が CF と同程度ユーザの「知っている・好みのアイテム」を推薦できていることが示されている。提案手法では、予測評価値の高いアイテムを優先しながらトピックを多様化しているため、トピック多様化による Accuracy の低下が抑制できたといえる。なお、有意差はないものの、提案手法において値が低い理由としては、提案手法では全てのトピックを網羅するように推薦リストを作成するために、アイテムの嗜好傾向の幅が広くなり、結果的に知っているアイテムが推薦される可能性が減ったことが原因と考えられる。

ユーザの「好みのアイテム」を推薦することに関しては、対応のある平均値の差の検

定結果より、提案手法が CF と同程度ユーザの「好みのアイテム」を推薦できていることが示されている。なお、提案手法と CF との間に有意な差は見られなかったが、提案手法の方が CF よりも高い結果を得ている。これは、図 4-5 におけるユーザの「知っている・好みのアイテム」の結果において、提案手法が CF よりも値が低かったことが原因となっているためと考えられる。

ユーザの「知っているアイテム」の推薦と、ユーザの「知らないアイテム」を推薦することはトレードオフの関係であるが、ユーザの「知っている・好みのアイテム」を推薦することと、ユーザの「知らなかったが、推薦されて興味を持ったアイテム」を推薦することは、必ずしもトレードオフの関係になるとは限らない。その点において、提案手法は、Accuracy を維持しながら Novelty を向上させることができていることから、CF よりも有効な推薦手法であるといえる。

以上より、推薦リスト内におけるトピックの多様化手法が、Novelty の高い推薦を実現する有効な推薦手法であることが実験より示された。ただし、現状の研究の制約として、対象が DVD アイテムのみであること、被験者が学生のみであること、比較アルゴリズムがユーザベースの協調フィルタリングのみであることがある。今後多くの異なる被験者と、様々なアイテム、アルゴリズムによって提案手法の有効性を検証する必要があるだろう。

4.6. 結論と今後の課題

本研究では、推薦の目新しさという観点に注目し、目新しいアイテムを推薦することでユーザ満足度を向上させる推薦手法の構築を目的とした。我々は、嗜好傾向が類似するアイテムを複数推薦するよりも、嗜好傾向の異なるアイテムを複数推薦の方が、ユーザの知らないアイテムを推薦する可能性が広がるという仮定にもとづき、推薦リストの嗜好傾向の多様化により推薦の目新しさ (Novelty) を改善させる手法を新たに提案した。具体的には、まず、ユーザのアイテム評価情報をもとに、共評価が行われているアイテム同士をエッジで繋いだネットワークを生成する。つぎに、これをクラスタリングすることにより、嗜好傾向の類似するアイテムをトピックとして分類し、そのトピックが複数に及ぶようにアイテムを選定することで嗜好傾向の多様化を実現する。評価実験として、Amazon.co.jp の DVD アイテムを用いて多様化を行った推薦リストをユーザに提示し、ユーザ満足度のアンケートを行った。実験結果より、提案手法が既存のユーザベースの協調フィルタリング手法よりも、Novelty の高い推薦、つまり、ユーザの「知らなかったが、推薦されて興味を持ったアイテム」をより多く推薦できることを確認し

た。

今後の課題として、ユーザベースの協調フィルタリング以外でもトピックの多様化が有効であるかを検証することがある。協調フィルタリング手法の欠点として、評価されるアイテムに偏りが生じる「評価値疎ら問題」 [Claypool 99] や、それにより運用の初期段階ではユーザの満足する推薦が得られないといった「cold-start 問題」 [Schein 02] があげられる。そこで、今後はアイテムベースの協調フィルタリングや、コンテンツベースフィルタリング、あるいは、他の多様化推薦手法 [Ziegler 05, 茂木 10] との比較もおして、提案手法の有効性を示す必要がある。さらに、正確さと目新しさ (Novelty) のトレードオフの問題についてより詳細な分析を行うことである。今回は、提案手法によりユーザの「知らなかったが、推薦されて興味を持ったアイテム」を多く推薦できたが、その分、推薦の正確さ (ユーザの「知っている・好みのアイテム」の推薦) を犠牲にしている。今後、推薦リスト全体への評価など、評価方法についても検討し、ユーザは推薦にどの程度の Novelty を求めるのか検証していくことが課題となる。

5章. 考察

5.1. 研究成果のまとめ

3章での成果

3章では、ユーザの明示的な情報提供によって情報共有が行われる Q&A サイトに対して、サイトの運用者が操作可能である外的モチベーションを支援する報酬制度に焦点をあて、効果的な情報蓄積のための報酬制度を議論するためのエージェントベースモデルを提案した。そして、提案したモデルに基づいてシミュレーション実験を行うことで、Q&A サイトにおける報酬制度が蓄積される情報の数と質にもたらす効果を分析した。

シミュレーション実験の結果、金銭的報酬制度は、得られる回答の質の向上の観点においては効果的であるが、得られる回答の数は減ってしまうという欠点があり、金銭的な報酬は必ずしも万能な支援とはいえないことを明らかにした。

本研究の成果として、Q&A サイトにおけるその効果に関する客観的な効果は明らかになっていないという現状において、効果的な報酬制度を分析・議論するためのモデルを提案し、その金銭的報酬制度の運用方針に関しては、サイトが重視する情報蓄積にそった運用方針が必要になるとの知見を示した。

4章での成果

4章では、ユーザの有用な情報の取得支援のために、受動的な情報取得を支援する推薦システムをとりあげ、課題とされている Novelty 向上による推薦の有用性の向上のために、課題に取り組んだ。ここでは、ユーザの労力を伴わず、適用範囲の広い方法で Novelty を向上させる手法の構築を目的に、ユーザの嗜好情報（コンテンツの評価情報）のみを用いて、コンテンツの評価の集中傾向にもとづいてトピックを抽出し、既存の CF のリストに対してトピックが多様になるようにコンテンツを再選定するという手法を提案した。

評価実験として、被験者アンケートとして既存の協調フィルタリング（CF）との比較を行い、その結果、提案手法は既存の CF に比べて Novelty が向上し、個々のコンテンツに対する評価の平均値が向上するという結果を得た。

結論として、ユーザの労力を伴わず、適用範囲の広い方法で、Novelty を向上させるソーシャルフィルタリングを実現した。また、個々のコンテンツに対する評価の平均値

が向上するという結果から、既存の CF に対して、提案する Novelty の向上手法を適用することで、全体的な推薦の有用性が向上するという知見を示した。

5.2. 研究の展望と限界

報酬制度に関して

本研究では、報酬に関して金銭的報酬を扱ったが、これ以外にも、ポイントや感謝の言葉など、金銭的報酬以外の報酬もあげられる。特に、これらの報酬は、現実には、報酬を支払うユーザのコストと報酬を受け取るユーザの利得は必ずしも一致しない場合も考えられる。例えば、ポイントや感謝の言葉などの報酬に関して、報酬を支払うユーザのコストに比べ、報酬を受け取るユーザにとってはその利得を高く見積もるといった場合もあげられる。本研究で提案したモデルの場合では、質問者と回答者のあいだでやり取りされる報酬の大きさは共通としていたが、異なる係数値を取り入れることで拡張可能であるといえる。

外発的モチベーション以外のモチベーションに関して

本研究では、外発的モチベーションである報酬に焦点をあてたモチベーション支援制度を扱ったが、外発的モチベーションのみの支援では、ユーザの主体的な行動のための内発的モチベーションを下げるアンダーマイニング効果といった問題を起こす可能性も考えられる。このような問題は、長期的な観点でのサイトの活性化を考える上で大きな問題となる。ただし、アンダーマイニング効果の逆に、はじめは外発的なモチベーションによって行動していても行動に楽しみを覚えることで次第に内発的モチベーションにもとづく行動に代わるエンハンシング効果といったものも期待できる。このため、より現実的なモチベーション支援の方針としては、例えば、情報共有サイトのモチベーション支援策としては、サイトに参加するユーザが少なく、ネットワーク外部性が働きにくい環境においては、報酬などの外発的モチベーションによって外部からのユーザの参加をうながし、多数のユーザが参加するようになったときには、外的報酬といった外発的モチベーションではなく、内発的モチベーションを支援するといったアプローチが有効であると考えられる。

このような観点に対し、本研究でのユーザのモデル化においては、ユーザのモチベーションの変化や、ネットワーク外部性の効果に関しては、今回の課題の対象外でありモデル化は行っていない。そのため、長期的な観点で情報共有サイトの活性化を考える上

では、これらのユーザ行動特性の変化によるサイトの発展の変化は重要な課題になってくる。今後、このような課題の設定や、そのためのモデル化に関しても検討していく必要があると考えられる。

有用な推薦に関して

本研究では、すべてのユーザに対しての Novelty 推薦を行ったが、その方法に関しては、ユーザのサイトへの利用状況に適した方法についても検討する必要がある。

本研究では、どの程度の Novelty の向上がユーザにとってより満足な推薦であるかといった調査は行っていない。ユーザにおける Novelty 推薦への期待は、個々のユーザの利用状況によって異なると考えられる。例えば、サイトを継続的に利用するユーザにとっては、推薦への飽きの可能性の観点から、より Novelty の高い推薦が必要となると考えられる。逆に、サイトを一時的に利用するようなユーザに対しては、Novelty の高い推薦よりも、そのときの情報取得目的に適した予測精度の高い推薦が必要となると考えられる。このため、今後個々のユーザの推薦システムに対する Novelty への期待の把握や、そのときの Novelty 向上の度合いなど、これを実現する推薦手法に関しても検討する必要があると考える。

情報蓄積と情報加工の連携に関して

本研究では、情報提供のモチベーション支援と、有用な情報推薦に関し、それぞれ個々の課題として扱っていたが、情報共有サイトにおいては、どちらの支援も重要な課題であり、これらはサイトに応じて適切な支援が必要になると考えられる。例えば、本研究では Q&A サイトにおいては、情報提供のためのモチベーション支援を扱ったが、サイトに参加する参加者が膨大になると質問者も増加し、それにより回答者が回答できる質問が埋もれてしまうといった問題もでてくることが考えられる。このような場合においては、回答者に対する情報提供を促すための支援（質問への接触機会の支援）も課題になると考えられるが、これに対して、本研究での Novelty 推薦は、回答者の持つ多様な知識の提供を促し、結果的に質問者に対する多様な回答を集めるうえで有効な方策の一つとして活用できると考えられる。

また、推薦を利用するサイトにおいても、サイトの運用初期における参加者数の不足や、参加者の特性の偏りによっては、他者情報にもとづく推薦に用いるコンテンツに対する評価情報などが集まらず、効果的な推薦が行えないといった問題も考えられる。こ

のような問題に対し、報酬といったインセンティブによって、評価情報などの推薦に用いる情報の蓄積を促すことで、効果的な推薦を実現するためにも有効な方法のひとつとして活用できると考えられる。ただし、これを検討するうえでも、モチベーション支援や推薦などがユーザにもたらす効果の理解と、そのためのユーザのモデルの構築、またその長期的な効果を分析するためのエージェントベースシミュレーションといったアプローチでの研究も必要になってくると考えられる。

6章. 結論と今後の課題

6.1. 結論

本研究の目的は、情報共有サイトの活性化を行うためのサイト設計に関する知見を明らかにすることである。そのために、本研究では、情報共有サイトにおけるユーザの参加モチベーションである「ユーザの有用な情報取得」を支援することが多数のユーザの継続的な参加を促し、サイトの活性化に繋がるうえで重要な要件ととらえ、これを支援するためのサイト設計を検討した。具体的には、情報共有サイトにおけるユーザの有用な情報取得支援の課題として、「(1) 情報提供のモチベーション支援」と、「(2) 有用な情報の取得支援」をあげ、これらの課題に対し有効な設計方策の知見を明らかにすることを目的とした。

「(1) 情報提供のモチベーション支援」に関しては、ユーザの明示的な情報提供によって情報共有が行われる Q&A サイトに対して、サイトの運用者が操作可能である外的モチベーションを支援する報酬制度に焦点をあて、効果的な情報蓄積のための報酬制度を議論するためのエージェントベースモデルを提案した。そして、提案したモデルに基づいてシミュレーション実験を行うことで、Q&A サイトにおける報酬制度が蓄積される情報の数と質にもたらす効果を分析した。シミュレーション実験の結果、金銭的報酬制度は、得られる回答の質の向上の観点においては効果的であるが、得られる回答の数は減ってしまうという欠点があり、金銭的な報酬は必ずしも万能な支援とはいえないとことを明らかにした。この研究課題の成果として、Q&A サイトにおけるその効果に関する客観的な効果は明らかになっていないという現状において、効果的な報酬制度を分析・議論するためのモデルを新たに提案し、金銭的報酬制度の運用方針に関しては、サイトが重視する情報蓄積にそった運用方針が必要になるとの知見を示した。この研究で得られた報酬制度のエージェントベースモデルと、そのシミュレーション結果の知見は、新たな Q&A サイトの構築や、既存の Q&A サイトにおける今後の運用方針を議論するうえでの判断材料になる知見として活用できるものである。

「(2) 有用な情報の取得支援」に関しては、受動的な情報取得を支援する推薦システムをとりあげ、課題とされている Novelty 向上による推薦の有用性の向上のために、課題に取り組んだ。ここでは、ユーザの労力を伴わず、適用範囲の広い方法で Novelty を向上させる手法の構築を目的に、ユーザの嗜好情報（コンテンツの評価情報）のみを

用いて、コンテンツの評価の集中傾向にもとづいてトピックを抽出し、既存の CF のリストに対してトピックが多様になるようにコンテンツを再選定するという手法を提案した。評価実験として、被験者アンケートとして既存の協調フィルタリング (CF) との比較を行い、その結果、提案手法は既存の CF に比べて Novelty が向上し、個々のコンテンツに対する評価の平均値が向上するという結果を得た。この研究の結論として、ユーザの労力を伴わず、適用範囲の広い方法で、Novelty を向上させるソーシャルフィルタリングを実現した。また、個々のコンテンツに対する評価の平均値が向上するという結果から、既存の CF に対して、提案する Novelty の向上手法を適用することで、全体的な推薦の有用性が向上するという知見を示した。

本研究の意義は、ユーザにとって有用な情報取得を支援する情報共有サイトの実現のために、有用な設計・運用方針の知見を示す点である。現在、ソーシャルメディアサイトは、多様に存在し、それらは実際に多数のユーザのもと運用されている。これらの多くは、ユーザの情報受発信を活発化させるために複数の機能や制度を導入しているが、どのような機能・制度がどう有効であったかという分析は、多数の機能・制度を導入している環境や、ユーザの多様なモチベーションのもとでは困難である。本研究は、従来の情報共有サイトにおける新たな機能や制度の導入における有効な方策を示し、ユーザにとって価値あるソーシャルメディアサイトの利活用に関して知見を加える研究である。

6.2 今後の課題

今後の課題として、ソーシャルメディアに関して、コミュニティといった観点での支援方策に関しても検討していく必要があるだろう。本論文では、Q&A サイトに対して個々のユーザのモチベーション支援として報酬を用いたが、高すぎる金銭的報酬は個々のユーザの援助的な動機にもとづく行動を失わせてしまう可能性があり [Pink 09]、このような報酬によるモチベーション支援は、内発的モチベーションを低下させる危険性もある。そのため、例えば、ユーザの援助的動機や貢献を促すための共同的なコミュニティの目的の設定や、ユーザのコミュニティへの貢献を評価する制度などによってモチベーションを高めるといった支援も必要になる。このような設計は、Wikipedia などの共同的なコミュニティとしての特徴をもったソーシャルメディアの支援としてより重要な課題になってくるだろう。

ソーシャルメディアにおいて、ユーザの情報発信は重要な要素であるが、ソーシャル

メディアをユーザ情報の蓄積とその提供に着目した捉え方ではなく、ユーザ間の双方向的なコミュニケーションを支援し、従来のメディアとの連携をはかることで、社会を支えるメディアとしての観点での支援も重要であると考えられる。

本研究では、価値あるソーシャルメディアを、ユーザに対して価値ある情報を提供するメディアとして捉えたが、ソーシャルメディアの価値はサイトに参加するユーザへの価値だけでなく、現実のコミュニティでの絆を深めるといった価値を持っている [総務省 10a]。そのため、価値あるソーシャルメディアの設計に関しては、ユーザのコミュニケーション支援し、社会を支えるメディアといった観点での支援方策も今後重要な課題になってくるといえるだろう。

謝辞

本論文をまとめるにあたり、指導教官であられる太田敏澄教授、審査委員であられる藤村考教授、大須賀昭彦教授、田野俊一教授、本多弘樹教授には、多くの貴重なご意見・ご指摘を頂きましたことを感謝いたします。

特に、指導教官であられる太田敏澄教授には、私が修士課程に入学してから博士課程を修了するまでの5年間、研究室生活を通して多大なご指導とご助言を頂きました。先生の幅広い知見や視野の広さにはいつも勉強させられてばかりで、研究の仕方や視点、特に社会情報という研究視点について、その重要性・おもしろさについて教えて頂き、今の私の研究の基本的な考え方になっております。この5年間、先生のご指導のもとで研究できましたことを心より感謝しております。

さらに、同研究室の諏訪博彦先生、同研究室 OB の山本仁志先生、岡田勇先生には、修士課程のころより研究へのご助言とご支援を頂きましたことを感謝申し上げます。諏訪先生には、公私ともに大変お世話になり、後期課程に入って思うように研究が進まないときでも的確なご指摘とあたたかいご助言を頂き、その後の研究の励みになりました。また、後輩との様々な研究議論の機会を作っていただきましたことを感謝しております。山本先生・岡田先生には、研究議論や様々な場をとおして、何よりも研究することの楽しさについて気づかせて頂きました。先生方の研究に対するスタンスは、私の今後の研究者としての目標ともなっています。

また、社会情報システム学講座の、藤村考教授、関良明准教授、山本佳世子准教授、また、M2、M1 の皆さまには、ゼミや研究審査を通してご指摘やご助言を頂きましたことを感謝いたします。さらに、今は修了した OB の皆さま、皆さんと一緒に研究生生活を過ごせたことは今でも楽しい思い出であり、この研究室で皆さんと出会えたことには本当に感謝しています。

そして最後に、これらの多くの人達との支援や出会いの期間を、最後まで陰ながら支えてくれた家族に、心より感謝申し上げます。

参考文献

- [Adomavicius 05] Adomavicius, G. and Tuzhilin, A.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, Vol.17, No.6, pp.734-749 (2005)
- [青山 08] 青山浩二, 鵜飼孝典, 小幡明彦, 原田裕明: 知識共有を動機付する手法, 人工知能学会 第3回知識流通ネットワーク研究会 (2008)
- [安藤 06] 安藤潤, 吉井伸一郎: WWW ナビゲーション向けコミュニティ分割手法に関する一考察, 情報処理学会研究会報告, Vol.2006, No.2, pp.115-122 (2006)
- [Axelrod 97] Axelrod, R.: The Dissemination of Culture: A Model with Local Convergence and Global Polarization, Journal of Conflict Resolution 41}, pp.203-226 (1997)
- [Balabanovic 97] Balabanovic M. and Shoham, Y.: Fab: Content-based, collaborative recommendation, Comm. ACM, Vol.40, No.3, pp.66-72 (1997)
- [Claypool 99] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M.: Combining Content-Based and Collaborative Filters in an Online Newspaper, Proc. of ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation (1999)
- [富士通総研 08] 富士通総研: ブログ・CGM 利用実態調査—商品選択時の情報源として重視される CGM—, 富士通総研 (2008)
- [Gilbert 99] Gilbert, G. N. and Troitzsch, K. G.: Simulation for the Social Scientist, Open University Press (1999)
- [Gilbert 99] Gilbert, G. N. and Troitzsch, K. G.: Simulation for the Social Scientist, Open University Press, (1999) (邦訳: ナイジェル・ギルバート, クラウス・G・トロイツェ (著), 伊庭崇, 高部陽平 (訳): 社会シミュレーションの技法 -政治・経済・社会をめぐる思考技術のフロンティア-, 日本評論社, 2003)
- [Herlocker 04] Herlocker, J., Konstan, J. and Riedl, J.: Evaluating collaborative filtering recommender systems, ACM Trans. Information Systems, Vol.22, No.1, pp.5-53 (2004)
- [石川 06] 石川徹也, 宇田隆幸: 情報フィルタリングの利用システム: 情報推薦システム, 情報の科学と技術, Vol.56, No.10, pp.458-463 (2006)
- [伊藤 07] 伊藤史: 消費者発信型メディア CGM, 毎日コミュニケーションズ (2007)
- [神寫 06] 神寫敏弘: 推薦システム—情報過多時代をのりきる, 情報の科学と技術 Vol.56, No.10, pp.452-457 (2006)

- [神 07] 神 敏弘：推薦システムのアルゴリズム (1), 人工知能学会誌, Vol.22, No.6, pp.826-837 (2007)
- [神 08] 神 敏弘：推薦システムのアルゴリズム (2), 人工知能学会誌, Vol.23, No.1, pp.89-103 (2008)
- [Kaplan 10] Kaplan, A.M. and Haenlein, M.: Users of the world, unite! The challenges and opportunities of social media, *Business Horizons*, Vol.53, No.1 (2010)
- [金谷 08] 金谷泰宏, 出口弘, 齋藤知也, 兼田敏之, 小山友介, 市川学, 田沼英樹：新型インフルエンザに対するパンデミック対策プログラム（<特集>エージェントベース社会シミュレーションの動向と展望）, *オペレーションズ・リサーチ：経営の科学*, Vol.53, No.12, pp.667-671 (2008)
- [岸田 03] 岸田和明：文書クラスタリングの技法：文献レビュー, *Library and Information Science*, No.49, pp.33-75 (2003)
- [Koch 04] Koch, S.: Profiling an open source project ecology and its programmers, *Electronic Market*, Vol.14, No.2, pp.77-88 (2004)
- [倉橋 01] 倉橋節也, 寺野隆雄, エージェントシミュレーションによる共同分配規範モデル, *電子情報通信学会論文誌, J84-D-I(8)*, pp.1160-1168 (2001)
- [Lawrence 99] Lawrence, P., Sergey, B., Rajeev, M. and Terry, W.: The PageRank Citation Ranking: Bringing Order to the Web, Technical Report, Stanford Digital Library Technologies. Project, pp.39-41 (1999)
- [Leibenstein 50] Leibenstein, H.: Bandwagon, Snob and Veblen Effects in the Theory of Consumers' Demand (1950)
- [Lerman 07] Lerman, K.: Social Networks and Social Information Filtering on Digg, *Proceedings of Int. Conf. on Weblogs and Social Media* (2007)
- [Li 08] Li, C. and Bernoff, J.: *groundswell: winning in a world transformed by social technologies*, Harvard Business School Press. (伊東奈美子 (訳) グランズウェル：ソーシャルテクノロジーによる企業戦略, 翔泳社) (2008)
- [Lin 07] Lin, H., Fan, W., Wallace, L. and Zhang, Z.: An empirical study of web-based knowledge community success, in *proceedings of the 40th Annual Hawaii International Conference on System Sciences(HICSS'07)*, pp.178c, Jan. (2007)
- [Linden 03] Linden, G., Smith, B. and York, J.: Amazon.com Recommendations: Item-to-Item Collaborative Filtering, *IEEE Internet Computing*, Vol.7, No.1 (2003)
- [松尾 05] 松尾豊, 友部博教, 橋田浩一, 中島秀之, 石塚満：Web 上の情報からの人間関係ネ

- ネットワークの抽出,人工知能学会論文誌,Vol.20, No.1E, pp.46-56 (2005)
- [McNee 06] McNee, S.M., Riedl, J. and Konstan, J.: Being accurate is not Enough: How Accuracy Metrics have hurt Recommender Systems, A Work-In-Progress paper in the Extended Abstracts of the 2006 ACM Conf.on Human Factor in Computing Systems (ACM SIGCHI), pp.1097-1101 (2006)
- [三浦 08] 三浦麻子, 川浦康至: 人はなぜ知識共有コミュニティに参加するのか: 質問行動と回答行動の分析,社会心理学研究 Vol.23, No.3, pp.233-245 (2008)
- [茂木 10] 茂木哲矢, 高村大也, 奥村学: ソーシャルタギングサービスを利用した多様性に基づく Web コンテンツ推薦, WebDB Forum 2010 (2010)
- [Mooney 99] Mooney, R.J. and Roy, L.: Content-Based Book Recommending Using Learning for Text Categorization, Proc. of ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation, (1999)
- [村上 09] 村上知子, 森紘一郎, 折原良平: 推薦の意外性向上のための手法とその評価人工知能学会論文誌 Vol.24(5), pp.428-436, (2009)
- [ネットレイティングス 08] ネットレイティングス株式会社: インターネット利用動向調査 (2008 年 11 月 デ ー タ) , http://csp.netratings.co.jp/nnr/PDF/Newsrelease12242008_J.pdf (2008)
- [Nexti 運営メンバー有志 10] Nexti 運営メンバー有志: NTT データ流ソーシャルテクノロジー, ソリューション IT 新書 (2010)
- [Newman 04a] Newman, M.E.J.: Fast algorithm for detecting community structure in networks, Phys. Rev. E, Vol.69, 066133 (2004)
- [Newman 04b] Newman, M.E.J.: Detecting community structure in networks, Eur. Phys. J. B38, pp.321-330 (2004)
- [Newman 04c] Newman, M.E.J.: Analysis of weighted networks, Phys. Rev. E70, 056131 (2004)
- [日本情報システム・ユーザ協会 01] 日本情報システム・ユーザ協会: 経営を変革するナレッジマネジメント~その研究と提言~, ナレッジマネジメント研究部会報告書, (2001).
- [NPO 法人ドットジェイピー 06] NPO 法人ドットジェイピー: Yahoo!みんなの政治, <http://seiji.yahoo.co.jp/> (2006)
- [Oded 07] Oded Nov: WHAT MOTIVATES WIKIPEDIANS ?, communications of the ACM, Vol.50, No.11, pp.60-64 (2007)

- [Ogawa 08] Ogawa.Y, Suwa.H, Yamamoto.H, Okada.I and Ohta.T.: Development of Recommender Systems Using User Preference Tendencies: An Algorithm for Diversifying Recommendation, Proc. of the 8th IFIP Conference on e-Business, e-Services, and e-Society (I3E 2008)), pp.61-73 (2008)
- [Ogawa 10] Ogawa.Y, Suwa.H, Yamamoto.H, Okada.I and Ohta.T.: Agent-Based Model of Q&A Community for Effective Pecuniary Payback System, Proc. of the 3rd World Congress on Social Simulation (WCSS2010) , (2010)
- [小川 09a] 小川祐樹, 諏訪博彦, 太田敏澄: 主体性の拡張を支援する推薦システムに関する研究, 2009 年日本社会情報学会 (JSIS&JASI) 合同研究大会 研究発表論文集, pp.372-375 (2009)
- [小川 09b] 小川祐樹, 諏訪博彦, 山本仁志, 岡田勇, 太田敏澄: 動的な動的なトピック分類に基づく Novelty を考慮した推薦アルゴリズムの提案, 情報処理学会論文誌, Vol.50, No.6, pp.1636-1648 (2009)
- [奥村 08] 奥村学, ブログマイニング技術の最新動向, 電子情報通信学会誌, No.91, Vol.12, pp.1054-1059 (2008)
- [大堀 08] 大堀耕太郎,高橋真吾: 社会シミュレーション技法におけるモデル評価の方法と課題,社会経済システム学会 (2008)
- [折田 09] 折田明子: 投稿を促すサイトの不思議 人はなぜ教えあうのか (2) インターネット - 最新ニュース :IT-PLUS, <http://it.nikkei.co.jp/internet/news/index.aspx?n=MMIT2E000003022009> (2009)
- [O'Reilly 05] O'Reilly, T: What Is Web 2.0: Design Pattern and Business Models for the Next Generation of Software, <http://oreilly.com/web2/archive/what-is-web-20.html> (2005)
- [Pazzani 96] Pazzani, M., Muramatsu, J. and Billsus, D.: Syskill and webert: Identifying interesting web sites, Proc. of Thirteenth National Conf. on Artificial Intelligence, pp.54-61, (1996)
- [Pink 09] Pink, D.H, Drive: The Surprising Truth About What Motivates Us, Riverhead Hardcover 1 edition (2009) (大前 研一 (訳) モチベーション 3. 0 持続する「やる気!」をいかに引き出すか, 角川書店 (2010))
- [Resnick 94] Resnick, P., Iacovou, N., Suchak, M., Bergstorm, P. and Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews, Proc. of ACM Conf. on Computer Supported Cooperative Work, pp.175-186 (1994)
- [Resnick 97] Resnick, P. and Varian, H.: Recommender systems, Comm. ACM, Vol.40, No.3,

- pp.56-58 (1997)
- [Sakaki 10] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors, Proceedings of WWW 2010, pp.851-860. (2010)
- [Sarwar 00] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Application of dimensionality reduction in recommender system, Proc. of ACM WebKDD Workshop (2000)
- [Sarwar 01] Sarwar, B., Karypis, G., Konstan, J. and Reidl, J.: Item-based collaborative filtering recommendation algorithms, Proc. 10th International World Wide Web Conf. (ACM), pp.285-296 (2001)
- [Schafer 01] Schafer, J., Konstan, J.A. and Riedl, J.: E - Commerce recommendation applications, Data Mining and Knowledge Discovery, Vol.5, pp.115 153 (2001)
- [Schein 02] Schein, A., Popescul, A., Ungar, L. and Pennock, D.: Methods and metrics for cold-start recommendations, Proc. 25th Annual ACM SIGIR Conf., pp.253-260 (2002)
- [Schelling 78] Schelling, T. C.: Dynamic Models of Segregation, Journal of Mathematical Sociology, Vol.1, pp.143-186 (1978)
- [清水 08] 清水拓也,土方嘉徳,西田正吾 : 発見性を考慮した協調フィルタリングアルゴリズム,電子情報通信学会論文誌,Vol.93-D,No.3, pp.538-550 (2008)
- [Smyth 01] Smyth, B. and McClave, P.: Similarity vs. Diversity, In proc. of the 4th International Conference on Case-Based Reasoning(ICCBR), pp.347-361 (2001)
- [総務省 10a] 総務省 : 平成 22 年版「情報通信白書」, <http://www.soumu.go.jp/johotsusintokei/whitepaper/h22.html> (2010)
- [総務省 10b] 総務省 : 平成 21 年「通信利用動向調査」, http://www.soumu.go.jp/main_content/000064217.pdf (2010)
- [Surowiecki 06] Surowiecki, J.: The Wisdom of Crowds, Broadway, NY: Broadway Books. (小高尚子 (訳)「みんなの意見」は案外正しい, 角川書店) (2006)
- [Swearingen 01] Swearingen, K. and Sinha, R.: Beyond Algorithms: An HCI Perspective on Recommender Systems, ACM SIGIR Workshop on Recommender Systems (2001)
- [庄司 08] 庄司昌彦: 地域 SNS サイトの実態把握 地域活性化の可能性, 情報通信政策研究プログラム研究成果論文 (2010)
- [庄司 10] 庄司昌彦: 「Twitter 政治」は民主主義を増進するか, 国際大学グローバル・コミュニケーション・センター, 智場 (115), pp.44-49 (2010)
- [Takahashi 07] Takahashi, S., Sallach, D., and Rouchier, J. (eds.): Advancing Social Simulation,

Springer (2007)

- [丹羽 06] 丹羽智史,土肥拓生,本位田真一: Folksonomy マイニングに基づく Web ページ推薦システム,情報処理学会論文誌,Vol.47,No.5,pp.1382-1392 (2006)
- [垂水 10] 垂水浩幸: ソーシャルメディアと実世界, 情報処理, Vol.51, No.7, pp.782-788 (2010)
- [寺野 04] 寺野隆雄: エージェント・ベース・モデリングへの招待,オペレーションズ・リサーチ・経営の科学,No.49, Vol.3, pp.131-136 (2004)
- [地方自治情報センター 07] 財団法人地方自治情報センター: 地域 SNS の活用状況等に関する調査, <http://www.lasdec.nippon-net.ne.jp/cms/resources/content/3686/result.pdf> (2007)
- [戸田 07] 戸田浩之,北川博之,藤村考,片岡良治: グラフ分析を利用した文書集合からの話題構造マイニング,電子情報通信学会論文誌,Vol.90-D, No.2, pp.292-310 (2007)
- [ダコプル 10] トニー・ダコプル, アンジェラ・ウー: ウィキを支えた無償投稿カルチャーの落日, <http://www.newsweekjapan.jp/stories/business/2010/09/post-1662.php> , 2010.9.8, (2010)
- [鳥海 08] 鳥海不二夫,石田健,石井健一郎: 小規模 SNS のモデル化と活性化シミュレーション,電子情報通信学会論文誌,B Vol.J91-B No.4 pp.397-406 (2008)
- [鳥山 09] 鳥山正博,菊池剛正,山田隆志,寺野隆雄: エージェントシミュレーションを用いた組織構造最適化の研究—スキーマ認識モデル—,電子情報通信学会論文誌 D, Vol.J92-D, No.11 ソフトウェアエージェントとその応用論文特集, pp.1919-1926 (2009)
- [梅田 06] 梅田望夫: ウェブ進化論, 筑摩書房 (2006)
- [Wu 07] Wu, J., Holt, C. R. and Hassan, A. E.: Empirical Evidence for SOC Dynamics in Software Evolution, The 23rd IEEE International Conference on Software Maintenance} (2007)
- [山田 08] 山田和明,中小路久美代,山本恭裕: オンラインコミュニティにおけるインセンティブメカニズムのモデル化,合同エージェントワークショップ&シンポジウム 2008 (2008)
- [山本 03] 山本仁志,石田和成,太田敏澄: 消費者間オンライン取引における評判管理システムの分析,経営情報学会誌,Vol.12, No.3, pp.55-69 (2003)
- [山本 09] 山本晶: Web マーケティング, 人工知能学会誌, Vol.24, No.4, pp.486-493 (2009)

- [山下 05] 山下清美, 川浦康至, 川上善郎, 三浦麻子: ウェブログの心理学, NTT 出版 (2005)
- [Yoshikawa 09] Yoshikawa, H., Yamaguchi, Y. and Nakamura, T.: Social Media in Japan, <http://www.briansolis.com>
- [湯田 06] 湯田聰夫, 小野直亮, 藤原義久: ソーシャル・ネットワーキングサービスにおける人的ネットワークの構造, 情報処理学会論文誌, Vol.47, No.3, pp.865-874 (2006)
- [Ziegler 05] Ziegler, C., NcNee, S.M., Konstan, J.A. and Lausen, G.: Improving Recommendation Lists through Topic Diversification, Proc. of WWW2005, pp.22-32 (2005)

関連論文の印刷公表の方法及び時期

- (1). 全著者名 : Yuki Ogawa, Hirohiko Suwa, Hitoshi Yamamoto, Isamu Okada, Toshizumi Ohta
論文題目「Agent-Based Model of Q&A Community for Effective Pecuniary Payback System」
Oct. 2010, Proc. of the 3rd World Congress on Social Simulation (WCSS2010)
(第 3 章の内容)

- (2). 全著者名 : Yuki Ogawa, Hitoshi Yamamoto, Isamu Okada, Hirohiko Suwa, Toshizumi Ohta
論文題目「Development of Recommender Systems Using User Preference Tendencies: An Algorithm for Diversifying Recommendation」
Oct. 2008, Proc. of the 8th IFIP Conference on e-Business, e-Services, and e-Society (I3E 2008), pp.61-73 (第 4 章の内容)

- (3). 全著者名 : 小川祐樹, 諏訪博彦, 山本仁志, 岡田勇, 太田敏澄
論文題目「動的な動的なトピック分類に基づく Novelty を考慮した推薦アルゴリズムの提案」
平成 21 年 6 月 情報処理学会論文誌 第 50 巻 6 号 (第 4 章の内容)

著者略歴

平成 18 年 3 月 大分大学工学部卒業

平成 18 年 4 月 電気通信大学大学院情報システム学研究科情報システム運用学専攻修士課程入学

平成 20 年 3 月 同上修了

平成 20 年 4 月 電気通信大学大学院情報システム学研究科社会知能情報学専攻博士課程入学

平成 23 年 3 月 同上修了見込