Master Thesis

Incentive Design for Soliciting Contributions to Q&A Communities

Supervisor Associate Professor Shigeo Matsubara

Department of Social Informatics
Graduate School of Informatics
Kyoto University

Asuka Shimada

February 12, 2016

Incentive Design for Soliciting Contributions to Q&A Communities

Asuka Shimada

Abstract

In recent years, the online Q&A services are becoming pervasive. Increasing numbers of people use these services as online sources for their information needs. Among these services, Q&A communities, also called "question-answer sites" or "answer sites", are growing in popularity with a number of users. Some of these websites employ some form of virtual rewards based on the contributions of user to motivate their users to actively participate on the site. These rewards, meant to provide an incentive for participation and contribution on the site, usually reflect various site-level accomplishments based on a user's cumulative 'performance' over life-long contributions.

This study aims to propose an incentive design for soliciting contributions to Q&A communities where new entrants appear sequentially.

In Q&A community, the competition of virtual rewards is regarded as a kind of contests. Contests are situations in which agents spend resources in order to win one or more prizes. A major feature that differentiates these types of interactions from the standard is that all contestants should bear the cost of their "bits". By regarding contributions to a Q&A community as bits, this problem can be viewed as a contest problem. However, it is difficult to model a Q&A community within the framework of existing contest theories, because new entrants are not considered in existing contest theories.

This study focuses following problems.

Modeling of the user behavior in Q&A communities considering new entrants

There are unspecified new entrants in a Q&A community. The existing users who have participated the Q&A community have some reputation points or badges according to their past contributions. Thus, they have different decision-making processes from new entrants who do not have any reputation points or badges. If there are no new entrants, this situation could be modeled by existing contest theories. However, it is difficult to model the situation where there are

some new entrants by existing theories. Hence, the new model that is considered the new entrants is required.

Analyzing the effect of incentive design on Q&A communities

It is necessary to evaluate the proposed models in order to analyze the effect of reward design on the users' behavior. Moreover, it is also necessary to analyze what kind of reward design encourage the participation of new users and contributions of all users.

To solve the first problem, I propose two models for Q&A communities by extending existing contest models. The first model is a model that rewards are allocated in order of the amount of contributions. The second model is a model that a reward is stochastically allocated according to the amount of contributions. Next, I regard a Q&A community to compete for virtual rewards as a two-stage contest and extended above two models by assuming that there are some users who participate the contest from the second stage. Also, I formulate optimal contributions for existing users and new entrants in the situation where contributions of existing users from first stages are carried over the second stage.

To solve the second problem, I conduct experiments by computer simulation because it is difficult to calculate optimal contributions mathematically. Finally, I analyze the effect of the incentive design on the Q&A community.

Contributions of this study are listed as follows:

Providing the model for Q&A communities considering new entrants

I construct two models for Q&A communities by extending the existing contest theories. I regard a Q&A community as a two-stage contest and I formulate the optimal contribution functions for existing users and new entrants

Clarifying the effect of incentive design on Q&A community

To analyze the effect of incentive design on the two-stage contest, I conduct some experiments based on simulation. As the result, the experimental results indicate the following design guidelines for soliciting contributions to Q&A communities.

- Better not to carry over reputation points (reputation points should be reset in a fixed time)
- Better to set larger rewards in later than earlier.

Q&A コミュニティへの貢献を引き出すインセンティブ設計

嶋田 飛鳥

内容梗概

近年，Q&A サイトが普及し，多くの人々がオンライン上の情報収集の手段として利用している．Q&A サイトの中でも，ユーザー同士の質問・回答で成り立つ Q&A コミュニティには，多くの利用者が存在する．

多くの Q&A コミュニティでは，ユーザーに継続的な参加を動機づけ，コミュニティを活性化させるために，バッジと呼ばれる仮想的な報酬を導入している．バッジは一度きりではなく，ユーザーの継続的な貢献に応じて様々な種類が用意されている場合がほとんどである．

本研究では，断続的に新規参入者が現れる Q&A コミュニティを想定し，Q&A コミュニティに適したインセンティブ設計を提案することを目標とした．

バッジを争う Q&A コミュニティは，コンテストの一種とみなすことができる．コンテストとは，複数の参加者が一つまたは複数の賞与を争って，各自の資源を投資するような状況をいう．参加者が「入札」として投資した資源は，コンテストの結果に関わらず浪費される．Q&A コミュニティへの貢献量を入札額とすることで，コンテストの問題に帰着できる．しかし，一般的なコンテストの理論では，コンテストへの新規参入を許していない．そのため，新規参入者がいるコミュニティのモデル化を既存のコンテストの議論の枠組みで行うことは難しい．

本稿の取り組む課題として，以下の二点である．

新規参入者がいる Q&A コミュニティに適したモデル構築

Q&A コミュニティでは，断続的に新規ユーザーが出現する．既に Q&A コミュニティに参加している既存ユーザーは，新規ユーザーと違い，過去の貢献に応じたバッジを持っているため新規ユーザーとは違った意思決定をすると考えられる．新規参入を許さない状況であれば，既存のコンテスト理論の枠組みで議論が可能であると考えられるが，新規参入者のいる Q&A コミュニティは既存のコンテスト理論でモデル構築を行うことが難しい．そのため，新規参入を考慮したモデルを提案する必要がある．

インセンティブ設計がユーザーに与える影響の分析

提案したモデルにおいて，バッジ等のインセンティブがユーザーやコミュニ

ティに与える影響を分析する必要がある．また，どのようなインセンティブ設計が，新規ユーザーの参加を促し，全体の貢献量を引き出すかを分析する必要がある．

まず第一の課題に対して，既存のコンテストのモデルを拡張し，新規参入者のいない場合のモデルを2種類構築した．1つ目のモデルは，費やした貢献量の順にバッジが分配されるモデル，2つ目のモデルは，貢献量に応じて確率的にバッジが分配されるモデルである．次にバッジを争うQ&Aコミュニティを二段階のコンテストとしてみなし，二段階目のコンテストから参加する新規ユーザーがいるものとして前述の2種類のモデルを拡張した．一段階目のコンテストから参加している既存ユーザーの貢献は，二段階目のコンテストにも引き継がれるものとすることで，既存ユーザー・新規ユーザーそれぞれの均衡戦略の導出を行った．

第二の課題に対しては，モデル上で数学的に確率や効用などを計算してユーザーの最適戦略を求めることが困難であるため，いくつかの状況のもとで，計算機によるシミュレーションを行い，既存ユーザー・新規ユーザーの最適貢献量を求めた．最後に得られた貢献量をもとに，インセンティブ設計が総貢献量にどのような影響を与えるかについて分析を行った．

本研究の貢献は次の2点である．

Q&Aコミュニティにおける新規参入者を考慮したモデル構築

既存の全員支払いオークションの理論を拡張し，Q&Aコミュニティのためのモデルを2種類構築した．Q&Aコミュニティへの貢献を二段階のコンテストと考えることで，既存ユーザー・新規ユーザーのそれぞれの均衡貢献量の定式化を行った．

インセンティブ設計がQ&Aコミュニティに与える影響の解明

二段階コンテストにおいて，報酬の設定や第一コンテストから参加している既存ユーザーの存在がQ&Aコミュニティ全体に及ぼす影響を分析するために，計算機シミュレーションを行った．また，得られた結果を分析することにより，

- 過去の貢献を引き継がないようにする（一定期間でリセットする）
- 後の段階に少し多め報酬を設定する

というQ&Aコミュニティを活性化させるためのインセンティブ設計の指針を示した．

Incentive Design for Soliciting Contributions to Q&A Communities

Contents

Chapter 1 Introduction

In recent years, the online Q&A services are becoming pervasive. Increasing numbers of people use these services as online sources for their information needs. Among these services, Q&A communities, also called "question-answer sites" or "answer sites", are growing in popularity with a number of users. Some of these websites employ some form of virtual rewards based on the contributions of user to motivate their users to actively participate on the site. These rewards, meant to provide an incentive for participation and contribution on the site usually reflect various site-level accomplishments based on a user's cumulative 'performance' over life-long contributions.

For example, the hugely popular Q&A community StackOverflow[1)] awards badges[2)] to users for various kinds and degrees of overall contributions, such as an archeologist badge for 'Edited 100posts that were in active for 6 months', or a Civic Duty badge for 'Voted 300 or more times'. These badges are large incentives for their users and StackOverflow are widely known as one of the most successful Q&A communities.

These virtual rewards, such as badges, motivate users to contribute the Q&A communities. Therefore, it is important for Q&A communities to design virtual rewards. The community designer might consider the designing rewards affects users' contributions and the number of users.

In another viewpoint, the Q&A communities, which employ virtual rewards such as reputation points or badges, are regarded as the contests to compete for these virtual rewards. Contests are situations in which agents spend resources in order to win one or more prizes. A major feature that differentiates these types of interactions from standard contests is that, independently of success, all contestants bear the cost of their "bids". By regarding contributions to Q&A a community as bits, this problem can be viewed as a contest problem. Numerous applications of such winner-take-all grand contests have been made to rent-seeking and lobbying in organizations, R&D races, political contests, promotions

1) The original Q&A community of the Stack Exchange network

2) `http://stackoverflow.com/help/badges`

in labor markets, trade wars, military and biological wars of attrition.

However, in the contests on Q&A communities such as StackOverflow, there are unspecified new entrants. In the huge Q&A communities such as StackOverflow, many new entrants come almost daily. These new entrants do not have any reputation points or badges obviously. Thus, they have different decision-making processes from existing users who have some reputation points or badges.

In this study, I aim to propose incentive design for Q&A communities to be able to consider the situation where there are some new entrants. As mentioned above, a Q&A community to compete for virtual rewards is regarded as a kind of a contest. However, it is difficult to model a Q&A community within the framework of existing contest theories, because new entrants are not considered in existing contest theories. To tackle this problem, I consider a Q&A community to compete for the virtual rewards as a two-stage contest and assumed that new entrants participate the contest from the second stage. Specifically, I assume the situation where some rewards are allocated in the order of contributions of users or where a reward is stochastically allocated according to contributions of users. Assumed that contributions of existing user in the first stage are carried over the second stage, I model the strategies of existing and new users. Next, I simulate the effect of designing rewards on the users' behavior when the users' ability follows the uniform distribution.

This study focuses on the following problems.

Modeling of the user behavior in Q&A communities considered new entrants There are unspecified new entrants in a Q&A community. The existing users who participate the Q&A community have some reputation points or badges according to their past contributions. Thus, they have different decision-making processes from new entrants who do not have any reputation points or badges. If there are no new enter users, these situations could be modeled by existing contest theories. However, it is difficult to model the situation where there are some new entrants by existing theories. Hence, the new model that is considered the new entrants is required.

Analyzing the effect of incentive design on Q&A communities I must evaluate the proposed models to analyze the effect of designing rewards on the users' behavior. Moreover, I must analyze that what kind of designing rewards encourage the participation of new users and contributions of all users.

The rest of this paper is organized as follows. Chapter 2 describes the online Q&A services. The online Q&A services can be divided three classes, the digital reference services, the expert services and the Q&A communities. I introduce these services respectively. Chapter 3 describes related works about contest theories and incentive design for Q&A communities. Chapter 4 and Chapter 5 propose and evaluate contest models for Q&A communities considering new entrants respectively. The model of Chapter 4 is based on order statics and the model of Chapter 5 is based on rent-seeking theories. Chapter 6 discusses incentive design for Q&A communities by comparing these two models. Finally Chapter 7 concludes this paper.

Chapter 2 Online Q&A Services

In this section, I introduce the outline of the online Q&A services.

The online Q&A services can be divided into three classes - "Digital reference services", "Expert services", and "Q&A communities"[1]. Table 1 shows the descriptions and services examples of the three types of the online Q&A sites.

Section2.1 explains Digital reference services, section2.2 explains Expert services, and section2.3 explains Q&A communities respectively.

2.1 Digital Reference Services

Digital reference services are the extension of library traditional reference services, and enable library users to access such services more easily and conveniently through various Web applications. Traditional library reference services employ expert researchers to help people find useful information. Digital reference services take place in the form of one-to-one interactions between a refer-

Table 1: Classification of the online Q&A site

	Description	Example of services
Digital reference services	The online analogue to library reference services[2]	AskERIC, Internet Public Library
Expert services	Question asking and answering services offered by various types of commercial and noncommercial organizations other than libraries.	NetWellness, AllExperts etc.
Q&A communities	Community-based and purposefully design to support people who desire to ask and answer questions.	Yahoo!Answer, Stack-Overflow etc.

ence librarian and a user, and the overall process is usually confidential. Today, many public libraries have added such as AskERIC[1], Internet Public Library[2], and etc.

2.2 Expert Services

Expert services are question-asking and answering services offered by various types of commercial and non-commercial organizations other than libraries. These services are staffed by "experts" (of varying credentials), often in a relatively circumscribed topic area, such as science or oceanography.

Expert services can be divided into two categories: free services and fee services. Government agencies and non-profit organizations have developed free services responding to questions on the Web; examples include NetWellness[3]. Fee-based expert services are more likely to follow the person-to-person information consulting model. Service organizations provide gateways to experts in topic areas. Answerers are supposed to be subject experts, and their qualifications vary depending on the organization providing the service. Subject experts might have appropriate academic degrees or professional training. These services receive applications and examine the qualification of candidates, but the selection criteria or standards are not known. AllExperts[4] and JustAnswer[5] are examples of expert services available on Web.

2.3 Q&A Communities

Q&A communities are community-based and purposefully designed to support people who desire to ask and answer questions, interacting with one another online. Anyone can ask and answer questions, and evaluate questions or other users each other. This evaluation by other users is an incentive to participate the communities. There are many Q&A communities which reward users for their continuous contributions in the form of reputation points or badges. Many

1) `http://eric.ed.gov/`
2) `http://www.ipl.org/`
3) `http://www.netwellness.org/`
4) `http://www.allexperts.com/`
5) `http://www.justanswer.com`

Table 2: Several of the more popular Q&A communities.

Q&A community	URL	Description
StackOverflow	`http://stackoverflow.com/`	Q&A community for programmer
Yahoo!Answers	`http://answers.yahoo.com/`	Most popular Q&A community
Knowledge-iN	`http://kin.naver.com/`	Also called Ji-Sik-In, first Q&A communities, Korean language.
Quora	`http://www.quora.com/`	Hybrid Q&A community/wiki format
Baidu Zhidao	`http://zhidao.baidu.com/`	Also called Baidu Knows, Chines language

Q&A communities do not provide qualified experts, though they offer a function to identify contributors as peer experts, given users' assessments of their past contributions in particular topic areas.

Shah et al. proposed for Q&A communities as a site or service requiring[3]:

- A method for users to present an information need in the form of a natural language question(as opposed to a keyword query)
- A forum for public response
- A community, based on participation level, in which the above transactions are embedded.

Table2 shows several of the more popular Q&A communities.

2.3.1 StackOverflow

StackOverflow is known as the most successful Q&A community. As of January 2015, StackOverflow, the original Q&A community site of Stack Exchange network, has accumulated over 3.8 million registered users, who have asked 8.6 million questions and have provided 15 million answers[4]. This corpus has been built since 2008; the community receives 7 million visits a day with an average of 7,300 questions per day. In total 75% of all questions have an answer and

57% of questions have an answer that the questioner accepted as “best answer”.

StackOverflow is managed by a reputation system that is a kind of user-driven reputation system[5]. The users can explicitly vote on and give or take away points to producers of questions and answers. These reputation points are important as they give a user privileges that allow the user to interact with and manage community.

Asking and answering

On StackOverflow, general rules require questions to be “specific to computer programming or the computer programming or the computer programming profession”. Questions are also required to be answerable, well researched, and useful for many users. These restrictions on StackOverflow are something that makes it an early success in Q&A community[6].

There are many means of interaction on StackOverflow. Once a question is asked, it can be edited by users of the community with that privilege, commented on by the community, and voted on by eligible users. A question could reply many excellent competing answers, and some could receive hundreds of votes. On the other hand, a question could receive many down votes and be closed from receiving further answers. Some of these questions will also be deleted form the corpus, which takes a collaborative effort from the community[7]. Figure 1 shows a screenshot of StackOverflow's top page.

Badges

The users could get a “badge” as a symbol of contribution by the founders when they fulfill certain conditions. Badges are listed on their profile pages and other users can access them freely. Badges are divided according to the difficulty of getting them such as: “Gold”, “Silver”, and “Bronze”, and can be classified by each category (Table 3).

Reputation system

StackOverflow reputation system was built by the founders of the system to be a “rough measurement of how much the community trusts you; it is earned by convincing your peers that you know what you're talking about. ” The full ways to earn points is described in Table 4

Having 10 points is needed for users to vote up on questions. Voted up

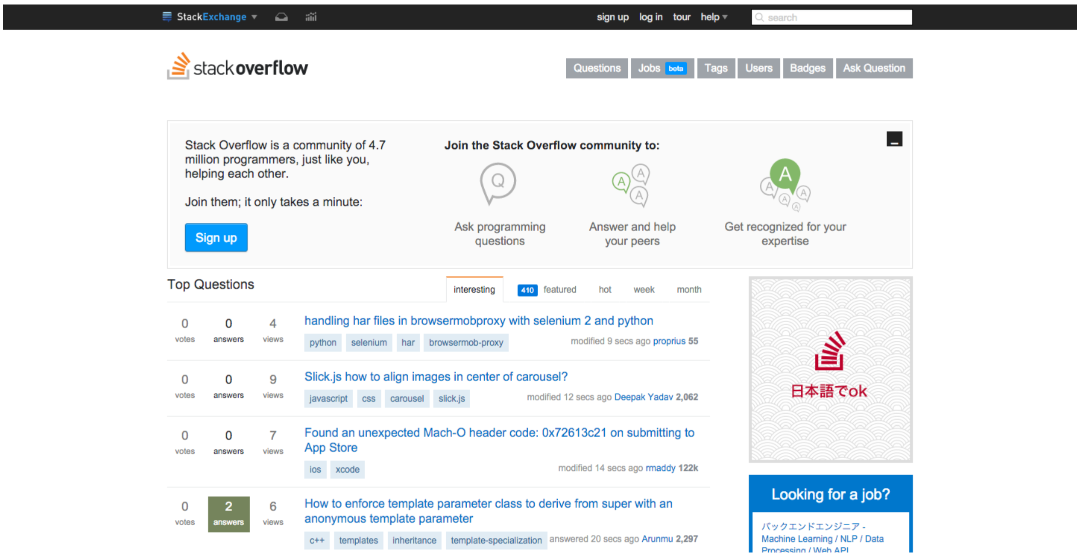


Figure 1: A screenshot of StackOverflow's top page

Table 3: The categories of badges in StackOverflow

Category	Description
Question Badges	Badges for question
Answer Badges	Badges for answer
Participation Badges	Badges for users' participation
Tag Badges	Badges for wiki score
Moderation Badges	Badges for moderation
Other Badges	Badges for other activities

is the greatest source of reputation points for all users. If the community become larger, the bias of the distribution of reputation points increases. Reputation points allow users access to privileges in the site. For example, having 15 points allows a user to vote a question or answer up. Table 5 shows a list of some of the major privileges that users can earn. The final privilege can be earned with 20,000 points. This table implies that about 76% users do not have vote up privilege. New users who do not have any points or privileges may be difficult to have the incentive to contribute.

Table 4: Earning Points on StackOverflow

Action	Result
Answer voted up	+10 points (Max +200 points a Day)
Question voted up	+5 points
Answer or question voted down	-2 points (-1 for a voter on questions)
Answer is accepted by questioner	+15 points (+2 to questioner)
Questioner offers bounty of own points for an accepted answer	Determined by questioner

Table 5: The example of the privileges on StackOverflow

Privilege	Points Required for Privilege	Percentage of Users with Privilege
Vote up	15 points	24%
Leave comments	5 points	12%
Vote down	125 points	8%
Vote to close any questions	3,000 points	1%
Vote to delete questions	10,000 points	0.1%

Chapter 3 Related Works

This Chapter describes the related works about this research.

I mainly focus the users' motivation and behavior and aim at proposing incentive design for Q&A communities. Q&A communities that employ the some form of virtual rewards can be modeled by the contest theories.

Section 3.1 describes the works about the contest theory, and section 3.2 describes the works about incentive design for Q&A communities.

3.1 Contest Theory

Many economic, political and social environments could be represented as contests in which competing agents have the opportunity to expend scarce resources - such as effort, money, time or troops - in order to affect the probabilities of winning prizes[8]. Examples range from the competition for mates, college admission, patents, research grants, or promotions within firms, to the process of litigation or lobbying politicians, to elections, sports competitions, and violent global conflicts.

The contest theories can be classified into three classes: the all-pay auction, the lottery contest, and the rank-order tournament. I introduce the all-pay auction and the lottery contest which are closely related to this research.

3.1.1 All-pay Auction

The all-pay auction is an auction in which every bidder must pay regardless of whether they win the prize, which is awarded to the highest bidder as in a conventional auction.

A Known theory about the all-pay auction theory is Baye's model[9]. Baye et al. characterized the entire set of equilibriums for the all-pay auction with complete information, a continuous strategy space, and possibly asymmetric prize valuations.

Moldovanu et al. discussed about optimal prize allocation for multi prize contest[10]. In their model, participants have privately informed a parameter (ability) affecting their costs of effort. The contestant with the highest effort has won the first prize and the contestant with second-highest effort has won the second prize. Moldovanu et al. modeled this contest as the all-pay auction. The contest's designer has maximized expected effort. Moldvanu et al. concluded that it have been optimal to allocate the entire prize sum to a single "first prize" when cost functions have been linear or concave in the effort.

I propose a model as an extension of Moldvanu's model to consider new entrants on Q&A communities in Chapter 4.

3.1.2 Lottery Contest

The Lottery contest is a contest in which every contestant must pay regardless of whether they win the prize, of which probability equals the ratio of a contestant's effort to the sum of each of the other contestants' efforts.

Rent-seeking contests are often modeled by the Lottery contests. In economics and public choice theory, rent-seeking involves seeking to increase one's share of existing wealth without creating new wealth. Examples of rent-seeking contests include Olympic bidding, corporate lobbying for a monopoly license or government contract and special interest lobbying for income redistribution programs.

Tullock developed what is perhaps the most widely used model of rent-seeking and recognized that because otherwise productive resources are used to compete for the rent, rent-seeking results in a social welfare loss[11][12]. Tullock's model of rent-seeking focuses on static, single-stage contests. The more an individual spends, the greater his chance of winning the prize. Specifically, in case that there are two persons, the probability of person A winning the prize is equal to $\frac{x_a}{x_a+x_b}$, where x_a and x_b are the expenditures of players A and B, respectively.

A large number of theoretical models emerged as extensions of Tullock's basic rent-seeking model. Anderson et al. presented an extended model assumed that a variable number of contestants, cost heterogeneity, and an entry fee. [13]. They predicted by their theoretical model that cost heterogeneity and an entry fee decrease participation.

Schmitt et al. presented an extended model in which a contestant's effort affects the probability of winning a contest in both the current and future stages in a repeated game context[14]. They reported that rent-seeking effort would be shifted forward from later to earlier stages, with no change in total rent-seeking expenditures about a non-carryover contest of the same duration.

I propose a model as extension model of Tullock's, Anderson's and Schmitt's model in Chapter 5. This model assumes that multi-stage contest, a variable number of contestants and cost heterogeneity.

3.2 Incentive Design for Q&A Communities

There are some researches about incentive design for Q&A communities. Easley et al. took a game-theoretic approach to badge design on Q&A communities, analyzing the incentives created by widely-used badge designs. In their model, winning a badge is valued, and effort is costly. Potential contributors to the site endogenously decide whether or not to participate, and how much total effort to put into their contributions to the site[15].

Immorlica et al. studied a model where agents seek badges on a website as a means of achieving social status within a Q&A community[16]. Designers can use these badges as a way to incentivize user contributions to their site. They found that mechanisms with a small number of badges or implement ranking of all agents were optimal within these settings.

However, the models of Easley and Immorlica have an issue that they do not consider the new entrants.

Chapter 4 Model based on Order Statics

In this chapter, I investigate the behavior of the users on Q&A communities by building a model based on the model which was proposed by Moldovanu et al. [10]. Moldovanu's model was based on order statistics[17]. In statics, the kth order statistics of a statistical sample is equal to its kth-smallest value.

Assumptions in this model are defined as follows.

Consider a website where q rewards are awarded. The value of the j-th reward is V_j, where $V_1 \geq V_2 \geq \ldots \geq V_q \geq 0$ The values of the rewards are common knowledge. I assume that $\sum_{i=1}^{q} V_i = 1$. This is just normalization.

The number of users is n. Each user has ability c_i. These are abilities to respond to the website, for example, the knowledge of programming in Stack-Overflow. Each user i makes a contribution x_i to this website. These contributions are submitted simultaneously. A contribution x_i causes a cost denoted by $c_i x_i$. $c_i > 0$ is an ability parameter. Note that a low c_i means that i has a high ability and vice-versa.

The ability (or type) of user i is private information to i. Abilities are drawn independently of each other from an interval $[m, 1]$ according to a distribution function F which is common knowledge. I assume that F has a continuous density $F' > 0$. To avoid infinite bids caused by zero cost, This research assumes that m, the type with the highest possible ability, is strictly positive.

The user with the highest contribution wins the first reward V_1. The user with the second highest wins the second reward V_2 and so on until all the rewards are allocated. The payoff of user i who has ability c_i and exert contribution x_i is either $V_j - c_i x_i$ if i wins reward j, or $-c_i x_i$ i does not win a reward.

Each user i chooses his contribution in order to maximize expected utility (given the other users' actions and the value of the rewards). I consider the utility for the designer that the designer maximizes the expected value of total contribution $E(\sum_{i=1}^{k} x_i)$.

I describe the case that no new entrants in section 4.1, and the case that new entrants exist in section 4.2. Moreover, I evaluate and analysis these cases in section 4.3.

4.1 A Case that No Entrants Exist

In this section, I explain the strategy of users in the case with no entrants.

I assume that all users undertake contribution according to the function b and that this function is strictly monotonic and differentiable. User i's maximization problem is:

$$max_x[(\sum_{k=1}^{q} V_{k\,n-1}C_{k-1}F(b^{-1}(x))^{k-1}(1-F(b^{-1}(x)))^{n-k})-ax] \tag{1}$$

${}_{n-1}C_{k-1}(F(b^{-1}(x)))^{k-1}(1-F(b^{-1}(x)))^{n-k}$ means the probability that user i with k-th contribution get the V_k reward. $(F(b^{-1}(x)))^{k-1}$ means the probability that there are $k-1$ users who have higher abilities than user i's. $(1-F(b^{-1}(x)))^{n-k}$ means the probability that there are $n-k$ users who have lower abilities than user i's. The probability that user i with k-th contribution get the V_k reward is ${}_{n-1}C_{k-1}(F(b^{-1}(x)))^{k-1}(1-F(b^{-1}(x)))^{n-k}$,because the cases with $k-1$ users who have higher abilities than user i's is ${}_{n-1}C_{k-1}$ varieties.

Let y denote the inverse of b. Using strict monotonicity and symmetry, the first-order condition (FOC) is:

$$\begin{aligned} 1 &= -\sum_{k=1}^{q} V_{k\,n-1}C_{k-1} \\ &\quad \times F(y)^{k-2}(1-F(y))^{n-k-1}y'\frac{1}{y}((n-1)F(y)-(k-1))F'(y) \end{aligned} \tag{2}$$

If the number of users is larger than the number of badges($q > n$), a user with the lowest possible ability $c = 1$ never win a reward. If the number of users is equal to the number of badges($q = n$), a user with the lowest possible ability $c = 1$ always win a $n-th$ reward. Hence, the optimal contribution of this type is always zero, which yields the boundary condition $y(0) = 1$.

Denote

$$\begin{aligned} G(y) &= \sum_{k=1}^{q} V_{k\,n-1}C_{k-1} \\ &\quad \times \int_a^1 \frac{1}{t}F(t)^{k-2}(1-F(t))^{n-k-1}[(n-1)F(t)-(k-1)]F'(t)dt \end{aligned} \tag{3}$$

The solution to the differential equation with the boundary condition is given by:

$$\int_y^0 dt = -G(y) \tag{4}$$

Equation (4) leads $x = G(y) = G(b^{-1}(x))$, and therefor that $b \equiv G$.

Thus, the contribution function of every user is given by:

$$\begin{aligned} b(c) &= \sum_{k=1}^{q} B_k(c)V_k \qquad (5) \\ B_k(c) &= {}_{n-1}C_{k-1}\int_c^1 \frac{1}{t}F(t)^{k-2}(1-F(t))^{n-k-1} \\ &\quad \times[(n-1)F(t)-(k-1)]F'(t)dt. \qquad (6) \end{aligned}$$

The equilibrium bid is equal to the sum of the weighted values of all rewards. These weighted values are different according to the expectation of getting rewards and users' abilities.

I have assumed that b is strictly monotonic and differentiable. Thus, I must prove that b is strictly monotonic and differentiable. To prove this, I prove the following lemma.

Lemma 1.

$$\sum_{k=1}^{q} B_k'(c) = -k_{n-1}C_k\frac{1}{c}F(c)^{k-1}(1-F(c))^{n-k-1}F'(c) \tag{7}$$

.

Proof. 1. *I prove this lemma by mathematical induction.*

Equation (6) leads

$$\begin{aligned} B_k'(c) &= -_{n-1}C_{k-1}[\frac{1}{c}F(c)^{k-2}(1-F(c))^{n-k-1}] \\ &\quad \times[((n-1)F(c)-(k-1)]F'(c) \qquad (8) \end{aligned}$$

(i) When $k=1$,

$$\begin{aligned} \sum_{k=0}^{q} B_1'(c) &= B_1'(c) \\ &= -(n-1)\frac{1}{c}(1-F(c))^{n-2}F'(c) \qquad (9) \end{aligned}$$

. Equation (8) is completed by Equation (9).

(ii) When $k = l+1$,if $k = l$ is completed,

$$\begin{aligned}
\sum_{k=1}^{q} B'_{l+1}(c) &= \sum_{k=1}^{q} B'_l(c) + B'_{l+1}(c) \\
&= -l_{n-1}C_l\frac{1}{c}(1-F(c))^{n-l-1}F(c)^{l-1}F'(c) \\
&\quad +_{n-1}C_l[\frac{1}{c}F(c)^{l-1}(1-F(c))^{n-l-2}] \\
&\quad \times[l-(n-1)F(c)]F'(c) \\
&= -_{n-1}C_l\frac{1}{c}F(c)^{l-1}(1-F(c))^{n-l-2} \\
&\quad \times[l(1-F(c))-(l-(n-1)F(c))] \\
&= -(n-l-1)_{n-1}C_l\frac{1}{c}F(c)^{l}(1-F(c))^{n-l-2}F'(c) \\
&= -(l+1)_{n-1}Cl+1\frac{1}{c}F(c)^{l}(1-F(c))^{n-l-2}F'(c) \\
&= \sum_{k=1}^{q} B'_{l+1}(c) \qquad (10)
\end{aligned}$$

Equation (8) is completed by Equation (10).

□

Lemma 1 and equation 7 leads:

$$\begin{aligned}
b'(c) &= \sum_{k=1}^{q} B'_k(c)V_k \\
&\leq V_k(\sum_{k=1}^{q} B'_k(c)) \\
&\leq 0 \qquad (11)
\end{aligned}$$

, where $V_1 \geq V_2 \geq \cdots V_q$ and $\forall c \in [m,1)$.

4.2 A Case that New Entrants Exist

Section 4.1 describes the strategies of users with no entrants. This section introduces the strategies of users with new entrants.

I regard the contest to compete for rewards as the two-stage contest to model the situation where existing users and new entratns are mixed. New entrants participate the contest in the second stage. I assume that the contributions of existing users in the first stage are carried over the second stage.

To simplify, I assume that there are one existing user and n_2 new entrants and q rewards in the second stage. In the first stage, n_1 users compete with one another. One of them (existing user) participates the second-stage. Equation (5) and (6) lead the contribution function b_1 of an existing user:

$$\begin{aligned} b_1(c) &= \sum_{k=1}^{q} B_k^1(c) V_k^1 \\ B_k^1(c) &= {}_{n_1-1}C_{k-1} \int_c^1 \frac{1}{t} F^1(t)^{k-2} (1 - F^1(t))^{n_1-k-1} \\ &\quad \times [(n_1 - 1) F^1(t) - (k-1)] F'^1(t) dt \end{aligned}$$

,where F^1 is an ability distribution function of first stage users, $F'^1 > 0$ is a continuous density of F^1 and V_k^1 is the values of rewards in the first stage. The existing user who has c_0 ability has invested $b_1(c_0)$ contribution in the first stage. These expenditures become sunk costs. In this regard, the decision of the existing user is different from of n_2 new entrants.

I assume new entrants decide their contributions by contribution function $b_2(c)$ when the number of the existing user is one. If x represents additional contribution of the existing user in the second stage, the existing user wins a new entrant who have c ability where $b_1(c_0) + x > b_2(c)$. Thus, the existing user's maximization problem is:

$$\begin{aligned} max_x[(\sum_{k=1}^{q} V_k^2 {}_{n_2-1}C_{k-1} F^2(b_2^{-1}(b_1(c_0)+x))^{k-1} \\ (1 - F^2(b_2^{-1}(b_1(c_0)+x)))^{n_2-k}) - c_0 x]. \end{aligned} \tag{12}$$

When x^* represents the optimal value of x, the new entrants think that the existing user has c_0' ability, where $c_0' = b_2^{-1}(b_1(c_0) + x^*)$. I call this c_0' ability as the appearance ability of the existing user.

Next, I discuss the contribution of the new entrant i who have c_i ability. I assume that the optimal contribution of the existing user at the second stage is x^*.

Case that $c_i < c_0'$,

The new entrant i's maximization problem is:

$$max_x[(\sum_{l=1}^{q} V_l^2 \sum_{k=l}^{n_2} {}_{n_2-1}C_{k-1}F^2(c_0')^{k-1}(1-F^2(c_0'))^{n_2-k}$$
$${}_{k-1}C_{l-1}F^2(b_2^{-1}(x))^{l-1}(1-F^2(b_2^{-1}(x)))^{k-l}) - cx] \tag{13}$$

,where F^2 is an ability distribution function of the second stage users, $F'^2 > 0$ is a continuous density of F^2 and V_l^2 is the values of rewards in the second stage. ${}_{n_2-1}C_{k-1}F^2(c_0')^{k-1}(1-F^2(c_0'))^{n_2-k}$ means the probability that there are k new entrants including i who have $c_i < c_0'$ abilities. $F^2(b_2^{-1}(x))^{l-1}(1-F(b_2^{-1}(x)))^{k-l}$ means the probability that new entrant i with $l-th$ contribution get V_l reward. The expected utility of the new entrant i who have $c_i < c_0'$ ability could be calculated by the sum of the expectation of getting each rewards whose value is $V_l^2(l = 1, 2 \ldots q)$.
Thus, the contribution function b_2 of enter users is given by:

$$b_2(c) = b_2(c_0') + \sum_{l=1}^{q} B_l^2(c)V_l^2 \tag{14}$$

$$B_l^2(c) = \sum_{k=l}^{n_2} {}_{n_2-1}C_{k-1\,k-1}C_{l-1}F^2(c_0')^{k-1}(1-F^2(c_0))^{n_2-k}$$
$$\times \int_c^{c_0'} \frac{1}{t}F^2(t)^{l-2}(1-F^2(t))^{k-l-1}$$
$$((k-1)F^2(y) - (l-1))F'^2(t)dt \tag{15}$$

$$c_0' = b_2^{-1}(b_1(c_0) + x^*) \tag{16}$$

Case that $c_i > c_0'$

The new entrant i's maximization problem is:

$$max[(\sum_{l=2}^{q} V_l^2 \sum_{k=0}^{l-2} {}_{n_2-1}C_kF(c_0')^k(1-F(c_0'))^{n_2-k-1} \tag{17}$$
$${}_{n_2-k}C_{l-k-2}F(b^{-1}(x))^{l-k-2}(1-F(b^{-1}(x)))^{n_2-l-2}) - cx]$$

The new entrant i is not able to get V_1^2 because his ability is lower than the existing user.
Thus, the contribution function b_2 of enter users is given by:

$$
\begin{aligned}
b_2(c) &= \sum_{l=2}^{q} B_l^2(c) V_l^2 & (18) \\
B_l^2(c) &= \sum_{k=0}^{l-2} {}_{n_2-1}C_{k\,n_2-k-1}C_{l-k-2} \\
& F^2(c_0')^k (1 - F^2(c_0'))^{n_2-k-1} \\
& \times \int_c^1 \frac{1}{t} F^2(t)^{l-k-3} (1 - F^2(t))^{n_2-l} \\
& [(n_2 + k + 3) F^2(y) - (l - k - 2)] F'^2(t) dt & (19)
\end{aligned}
$$

4.3 Evaluation of Model

In this section, I experiment and evaluate the model based on order statics. Section 4.1 describes the case that no entrant, and section 4.2 describes the case that new entrants exist. In these models, assumed that the ability's distribution, the number of rewards and the values of rewards are known by all users. The behavior of new entrants might change according to the ability and contribution of the existing user in the first stage. It is necessary to evaluate the effect of ability's distribution, the number of rewards and the values of rewards for users' contributions. Thus, I experiment this model by computer simulation.

4.3.1 The Method of Experiments

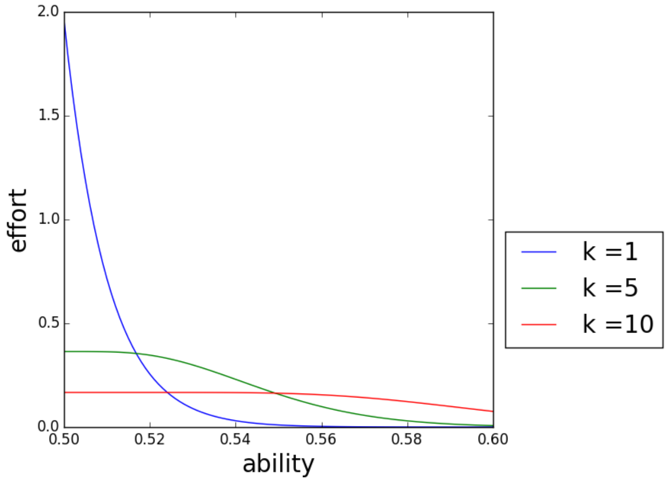
The method of experiments is as follows.

a)A case that no entrants exist

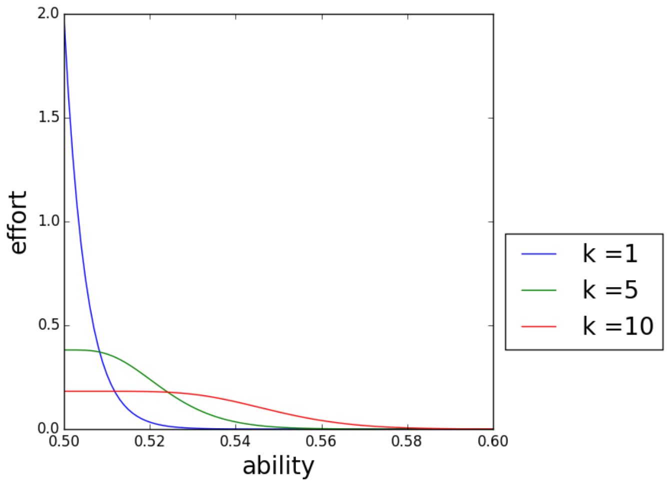
In this case, I examine the effect of the number of users, the number of rewards and the values of rewards on the users' contributions. The assumptions of this experiment are as follows: The number of users is set to either of 50 or 100, the number of rewards is set to either of 1, 5, 10. The abilities of users are according to uniform distribution between 0.5 and 1. The values of all rewards are equal and are normalized (that is $\Sigma_{i=1}^{p} V_i = 1$). For example, if the number of rewards is k, $V_i = \frac{1}{k}$.

b)A case that new entrants exist

In this case, I examine how the existing entrants affects the new entrants. The assumptions of this experiment are as follows: The number of users in the first stage is 50, the number of users in the second stage is 50, and the number



(a) A case that the number of user is 50



(b) A case that the number of user is 100

Figure 2: The amount of contributions for various rewards

of rewards is 5. The values of all rewards are equal and normalized. To compare the case that an existing user exerts large contribution in the first stage and the case that an existing user exerts small contribution in the first stage, I calculate the optimal contribution of new entrants when 5-th and 30-th user in the first stage participate the second stage, respectively.

4.3.2 Result of Experiments

Table 6: The amount of max contribution and total contributions.

(a) A case that the number of user is 50

The number of rewards	The amount of max contribution	The amount of total contributions
1	1.96	0.96
5	0.36	0.89
10	0.17	0.82

(b) A case that the number of user is 100

The number of rewards	The amount of max contribution	The amount of total contributions
1	1.98	0.98
5	0.38	0.94
10	0.18	0.90

A case that no entrants exist

Figure 2 shows the relation between abilities and the amount of contribution of users with the number of rewards (1,10,50). Figure 2a expresses the case that the number of users is 50 and Figure 2b expresses the case that the number of users is 100. I assume that the max ability of users is 0.5, and the minimum ability of users is 1(The lower the value of c becomes, the higher the ability becomes). Figure 2 indicates that the fewer the number of rewards is, the more contributions are solicited by users who have higher abilities and the higher the total amount of contributions is.

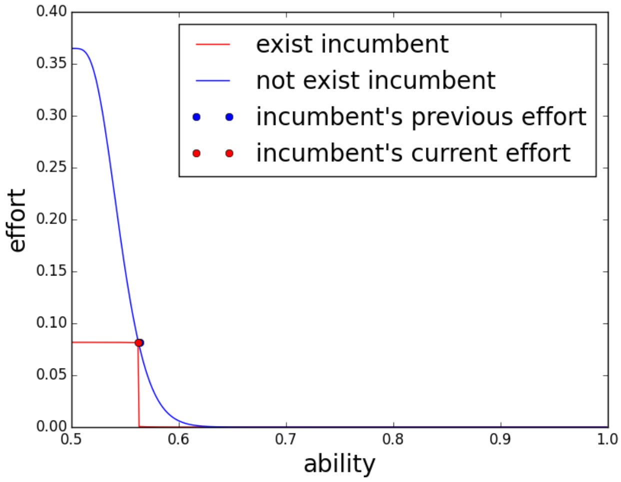
Table 6 shows the amount of max contribution and the total amount of contributions for each setting of rewards. The amount of max contribution means the amount of contribution in case of $ability = 0.5$ because the contribution function is a decreasing monotonic function. The total amount of contributions means $n \int_{0.5}^{1} b(c)dc$ assumed that the number of users is n.

This table indicates that if the number of users increases, the amount of max contribution increases. This is because if the number of users increases, the probability of getting rewards decreases. Thus, the users who have high ability should make more contributions in order to get rewards. Similarly, if the number of users increases, the total amount of contributions increases. This is because the number of users who make contributions hard increases.

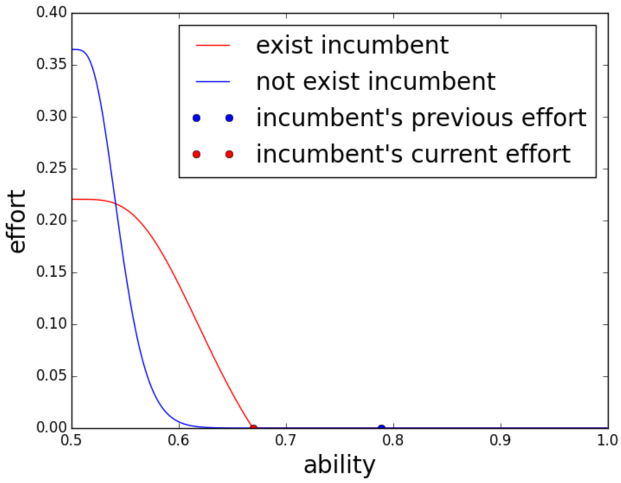
A case that new entrants exist

Figure 3 shows the amount of contributions of the second stage in the situation where the number of users in the first stage is 50, the number of users in the second stage is 50, the number of rewards is 5, and the number of existing user from the first stage is one. The blue line means the amount of contributions in case that no existing user exists, and the red line means the amount of contributions in case that an existing user exists. The blue point means the amount of contribution of the existing user in the first stage and the red point means the amount of contribution of the existing user in the second stage. Figure 3a shows the amount of contributions in case that an existing user who has been 5-th in the first stage and Figure 3b shows the amount of contributions in case that an existing user who has been30-th in the first stage.

I consider the case that an existing user who has been 5-th in the first stage exists. In this case, the amount of the contribution of the existing user in the first stage is about 0.8. However, his amount of the contribution in the second stage is almost zero. His appearance ability(c_0') is about 0.57, which is almost as much as his real ability. This is because his contribution of the first stage is carried over and the effect of cost on the contribution is greater than the increase of the expectation by the additional contribution of the second stage. The amount of contribution of the new entrants who have higher ability than the appearance ability of the existing user is slightly higher than the existing user. This is because there is an existing user whose probability of getting rewards. Thus, if new entrants who have higher ability than the existing user make an contribution slightly harder than the existing user, the probabilities of getting rewards become higher and the incentive to contribute becomes smaller. On the other hand, the amount of the contribution of the new entrants who have lower ability than the appearance ability of the existing user is almost zero. The incentive for new entrants who have lower ability than the existing



(a) A case that an existing user who has been 5-th in first stage exists.



(b) A case that an existing user who has been 30-th in first stage exists.

Figure 3: A case that an existing user exists.

user become smaller because there is an existing user who has higher ability and the probability that the new entrants get rewards become smaller.

Next, I consider the case that an existing user who has been 30-th in the first stage exists. In this case, the existing user makes a little contribution in the first stage and second stage. Just like a case that an existing user who

Table 7: The relation between the existence of existing user and the amount of the contributions

the order of existing user in first stage	max contribution	total contribution
no existing user	0.36	0.89
5	0.081	0.26
30	0.22	1.22

has been5-th in the first stage exists, this is because his contribution of the first stage is carried over and the effect of cost on the contribution is greater than the increase of the expectation by the additional contribution of the second stage. The appearance ability of the existing use becomes higher from 0.8 to 0.68. The new entrants who have lower abilities than the appearance ability of the existing user make little contributions. The new entrants who have higher abilities than the appearance ability of the existing user make higher contributions than the case that no entrants exist. This is because the probability that the existing user get the rewards is lower, the probability that the new users get the rewards becomes higher and the incentive to contribute become greater.

Table 7 shows the relation between the existence of existing user and the amount of contributions. In case that there is an existing user who has been5-th in first stage, the total amount of contributions is a third compared with a case that there are no existing user, but in case that there is an existing user who has been30-th in fist stage, the total amount of contributions is about 1.5 times.

I summarize the result of these experiments as follows:

- If the number of rewards decreases, the total amount of contributions increases.
- If the number of users increases, the total amount of contributions and the amount of max contribution increases.
- If there is an existing user who have high ability, the amount of contributions of new entrants decreases.
- If there is an existing user who have low ability, the total amount of contributions increases.

Chapter 5 Model based on Rent-seeking Theories

In this chapter, I propose a model of user behaviors on Q&A communities based on rent-seeking models([12], [13], [14]).

Assumptions in this model are defined as follows. Just like the model of Chapter 4, each user i makes a contribution x_i to this Q&A community. A contribution x_i causes a cost denoted by $c_i x_i$. $c_i > 0$ is an ability parameter. Note that a low c_i means that i has a high ability and vice-versa.

In the model of Chapter 4, the rewards are allocated in the order of the amount of contributions. In this model, a reward is stochastically allocated according to the amount of contributions. This differs from two models.

Assume that n risk-neutral users compete by expending contribution x_i ($i = 1, \ldots, n$) and that the probability that user i wins the contest is

$$P_i = \frac{x_i}{\sum_{k=1}^{n} x_k} \tag{20}$$

.

In this model, I assume that there is a reward V in a contest. Thus, the payoff of user i who has ability c_i and exert contribution x_i is either $V - c_i x_i$ if i wins a reward, or $-c_i x_i$ if i does not win a reward. Each user i chooses his contribution in order to maximize expected utility (given the other users' ability and the value of the reward).

I describe the case with no entrants in Section 5.1 and the case with new entrants Section 5.2.And I evaluate and analyze these cases in Section 5.3.

5.1 A Case that No Entrants Exist

This section introduces the strategies of users with no entrants.

Assume that n risk-neutral users compete for a reward V and each user i have ability c_i, the maximization problem of user i is:

$$max[\frac{x_i}{\sum_{k=1}^{n} x_k} V - c_i x_i]. \tag{21}$$

Let x_i^* denote the optimal contribution of user i, the first-oreder condition is:

$$\frac{\sum_{k=1}^{n} x_k - x_i^*}{(\sum_{k=1}^{n} x_k)^2} V - c_i = 0. \tag{22}$$

By adding together with $i = 0, \ldots, n$ in equation (22), I get:

$$\frac{(n-1)\sum_{k=1}^{n} x_k}{(\sum_{k=1}^{n} x_k)^2} V - \sum_{j=1}^{n} c_i = 0. \tag{23}$$

By solving equation (23) of an inverse $\sum_{k=1}^{n} x_k$, I get:

$$\sum_{k=1}^{n} x_k = \frac{n-1}{\sum_{k=1}^{n} c_k} V. \tag{24}$$

By substituting equation (24) for equation (22), I get:

$$x_i^* = \frac{V(n-1)}{\sum_{k=1}^{n} c_k}[1 - \frac{c_i(p-1)}{\sum_{k=1}^{n} c_k}]. \tag{25}$$

According to above equations, the optimal contribution x_i^* of user i could be formulated.

5.2 A Case with New Entrants Exist

This section describes the strategies of users with new entrants.

Just like section 4.2, I regard a contest to compete for a reward as a two-stage contest to model the situation where existing users and new entrants are mixed. New entrants participate the contest in the second stage. I assume that the contributions of existing users in the first stage are carried over the second stage. There are n existing users and m new entrants. Each user i's ability c_i is common knowledge.

In a repeated game context, contribution carryover implies that contribution made by user i in the first stage not only affect the probability user i will win the reward in the first stage but also the probability that user i will win a reward in the second stage. Therefore, the probability P_{i2} that the existing users who participate in the first stage win the reward in the second stage is a function of the existing users ’ contributions in the first and second stages and the new entrants ’ contributions in the second stage.

$$
\begin{aligned}
&Px_{i2} = f(x_{i1}, x_{i2}, x_{k1}, x_{k2}, y_{j2})\\
&for\ \forall k = 1, \ldots, n, k \neq i, j = 1, \ldots, m.
\end{aligned}
$$

x_{i1} means the first stage contribution of existing user i, x_{i2} means the second stage contribution of existing user i, y_{j2} means the second stage contribution of new enter use j.

Similarly the probability Py_{j2} that the new entrants win the reward in the second stage is represented as follows:

$$
\begin{aligned}
&Py_{j2} = f(x_{k1}, x_{k2}, y_j, y_l)\\
&for\ \forall k = 1, \ldots, n, l = 1, \ldots, m, j \neq l.
\end{aligned}
$$

A reasonable assumption is that total contributions carryover from the first stage to the second stage at a rate δ, where $0 \leq \delta \leq 1$. This is equivalent to assuming that contributions depreciate at a rate of $1 - \delta$ from the first stage to the second stage. Assuming that carryover is available for all existing users and that $\delta_k = \delta_l$ for all k and l, the game is symmetric. The probability Px_{i2} that existing user i wins the reward and Py_{j2} that new entrant j wins the reward in the second stage are:

$$Px_{i2} = \frac{x_{i2} + \delta x_{i1}}{\sum_{k=1}^{n}(x_{k2} + \delta x_{k1}) + \sum_{l=1}^{m} y_{l2}} \tag{26}$$

$$Py_{j2} = \frac{y_{j2}}{\sum_{k=1}^{n}(x_{k2} + \delta x_{k1}) + \sum_{l=1}^{m} y_{l2}}. \tag{27}$$

The contribution in each stage increases the probability of winning in the first and second stages. If $\delta = 0$, there is no carryover of contributions from first to the second stage.

Each existing user i chooses a contribution in the first stage to maximize his expected payoff over the second stage. Assuming the total value of the rewards of first and second stages is V, the value of the first stage reward is $(1-\alpha)V$ and the value of the second stage reward is αV, where $0 \leq \alpha \leq 1$. α means the distribution rate of reward for each stage. The individual's expected payoff in the first stage is:

$$\begin{aligned} Ex_{i2} = & \frac{x_{i1}}{\sum_{k=1}^{n} x_{k1}}(1-\alpha)V + \frac{x_{i2}+\delta x_{i1}}{\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2}}\alpha V \\ & - c_i(x_{i1}+x_{i2}) \end{aligned} \tag{28}$$

$$Ey_{j2} = \frac{y_{j2}}{\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2}}\alpha V - c_j y_{j2}. \tag{29}$$

Each existing user i and each new entrants j chooses a contribution to maximize his expected payoff. The first-order conditions are calculated by partial differentiations of equation(28) by x_{i1} and x_{i2} and equation(29) by y_{j2}. These first-order conditions are as follows:

$$\begin{aligned} \frac{\partial Ex_{i2}}{dx_{i1}} = & \frac{\sum_{k=1}^{n} x_{k1} - x_{i1}}{(\sum_{k=1}^{n} x_{k1})^2}(1-\alpha)V \\ & + \delta\frac{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2}) - (x_{i2}+\delta x_{i1})}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2})^2}\alpha V - c_i \\ = & 0 \end{aligned} \tag{30}$$

$$\begin{aligned} \frac{\partial Ex_{i2}}{dx_{i2}} = & \frac{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2}) - (x_{i2}+\delta x_{i1})}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2})^2}\alpha V - c_i \\ = & 0 \end{aligned} \tag{31}$$

$$\begin{aligned} \frac{\partial Ey_{j2}}{dy_{j2}} = & \frac{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m} y_{l2}) - (x_{i2}+\delta x_{i1})}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})^2}\alpha V - c_j \\ = & 0. \end{aligned} \tag{32}$$

By adding together with $i = 1, \ldots, n$ in equation (30) and (31) and with $j = 1, \ldots, m$ in equation(32), I get:

$$\frac{(n-1)\sum_{k=1}^{n}x_{k1}}{(\sum_{k=1}^{n}x_{k1})^2}(1-\alpha)V+\delta\frac{(n-1)\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+n\sum_{l=1}^{m}y_l}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m}y_l)^2}\alpha V=\sum_{k=1}^{n}c_k \tag{33}$$

$$\frac{(n-1)\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+n\sum_{l=1}^{m}y_l}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m}y_l)^2}\alpha V=\sum_{k=1}^{n}c_k \tag{34}$$

$$\frac{m\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+(m-1)\sum_{l=1}^{m}y_l}{(\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})+\sum_{l=1}^{m}y_l)^2}\alpha V=\sum_{l=1}^{m}c_l. \tag{35}$$

Equation (33),(34) and (35) lead:

$$\sum_{k=1}^{n}x_{k1}=\frac{(n-1)(1-\alpha)}{(1-\delta)\sum_{k=1}^{n}c_k}V \tag{36}$$

$$\sum_{k=1}^{n}(x_{k2}+\delta x_{k1})=\frac{\alpha V}{\sum_{l=1}^{m}c_l(1+\frac{\sum_{k=1}^{n}c_k}{\sum_{l=1}^{m}c_l})^2}(n-\frac{\sum_{k=1}^{n}c_k}{\sum_{l=1}^{m}c_l}(m-1))(m+n-1) \tag{37}$$

$$\sum_{l=1}^{m}y_l=\frac{\alpha V}{\sum_{l=1}^{m}c_l(1+\frac{\sum_{k=1}^{n}c_k}{\sum_{l=1}^{m}c_l})^2}(\frac{\sum_{k=1}^{n}c_k}{\sum_{l=1}^{m}c_l}m-(n-1))(m+n-1). \tag{38}$$

By substituting equation (36),(37) and (38) for equation(31) and (32) respectively, the optimal first stage contribution x_{i1}^* of existing user i, the optimal second stage contribution x_{i2}^* of existing user i and the optimal second stage contribution y_{j2}^* of new entrants j are formulated as :

$$x_{i1}^*=\frac{(n-1)(1-\alpha)V}{((1-\delta)\sum_{k=1}^{n}c_k)}(1-\frac{c_i(n-1)}{\sum_{k=1}^{n}c_k}) \tag{39}$$

$$x_{i2}^*=\frac{\alpha V(m+n-1)}{\sum_{k=1}^{n}c_k+\sum_{l=1}^{m}c_k}(1-\frac{c_i(m+n-1)}{\sum_{k=1}^{n}c_k+\sum_{l=1}^{m}c_l})-\delta x_{i1} \tag{40}$$

$$y_{j2}^*=\frac{\alpha V(m+n-1)}{\sum_{k=1}^{n}c_k+\sum_{l=1}^{m}c_l}(1-\frac{c_j(m+n-1)}{\sum_{k=1}^{n}c_k+\sum_{l=1}^{m}c_l}). \tag{41}$$

Assuming that:

$$x_{i1}^* > 0 \tag{42}$$

$$x_{i2}^* \geq 0 \tag{43}$$

$$y_{j2}^* > 0. \tag{44}$$

These mean that the contributions to the Q&A community are not negative. I assume:

$$\delta \leq \frac{1}{2} \tag{45}$$

to avoid the negative second stage contributions of existing users (equation (40)). Note that only the contribution for the first and second stage contribution depend on the rate of carryover δ. Thus, the first stage contribution x_{i1}^* of existing user i increases as the rate of carryover increases(given by equation (39)). However, x_{i1}^* increases as the second stage contribution x_{i2}^* of existing user i decreases(given by equation(40)) and the restriction of equation(43) do not satisfied.

In this case that x_{i2}^* is negative with a large α and a small δ and large x_{i1}^*, I assume $x_{i2}^* = 0$. Schmitt modeled the contest with no entrants as the multi-stage contest and assumed $\delta \leq \frac{1}{2}$, too.

Equation (41) and (44) leads:

$$c_j < \frac{\sum_{k=1}^{n} c_k}{n-1} \tag{46}$$

This equation is the condition for new entrants to participate the second stage.

Assumed that U is the total contribution of all users in first and second stages, equation (36),(37) and (38) lead:

$$
\begin{aligned}
U &= \sum_{k=1}^{n} x_{k1} + \sum_{k=1}^{n} x_{k2} + \sum_{l=1}^{m} y_{l2} \\
&= \sum_{k=1}^{n}(x_{k2} + \delta x_{k1}) + (1-\delta)\sum_{k=1}^{n} x_{k1} + \sum_{l=1}^{m} y_{l2} \\
&= \frac{m+n-1}{\sum_{k=1}^{n} c_k(\sum_{k=1}^{n} c_k + \sum_{l=1}^{m} c_l)}\alpha V + \frac{(n-1)(1-\alpha)}{\sum_{k=1}^{n} c_k}V \\
&= \frac{V}{\sum_{k=1}^{n} c_k + \sum_{l=1}^{m} c_l} \\
&\times ((n-1)(\sum_{k=1}^{n} c_k + \sum_{l=1}^{m} c_l) + \alpha(m\sum_{k=1}^{n} c_k - (n-1)\sum_{l=1}^{m} c_l)) \qquad (47)
\end{aligned}
$$

By adding together with $j = 1, \ldots, m$, equation (46) leads:

$$\sum_{l=1}^{m} c_l < \frac{m}{n-1}\sum_{k=1}^{n} c_k. \qquad (48)$$

Equation (48) leads:

$$m\sum_{k=1}^{n} c_k - (n-1)\sum_{l=1}^{m} c_l > 0 \qquad (49)$$

Equation (47) and (49) lead $U > 0$. With $\alpha = 1$, that is, when the value of reward of the first stage is 0 and the value of reward of the second stage is V, the total contribution U become the largest. Assumed that U_{max} is max of U, U_{max} is given by:

$$U_{max} = \frac{m+n-1}{\sum_{k=1}^{n} c_k + \sum_{l=1}^{m} c_l}. \qquad (50)$$

In this way, I could calculate the total contribution in case that the abilities of all users are known and $x_{i2}^{*} < 0$ is allowed.

5.3 Evaluation of Model

In this section, I experiment and evaluate the model based on the rent-seeking model. Section 5.1 describes the case that no entrant, and section 5.2 describes the case that new entrants exist. In this model, assumed that the abilities of users and every user knows the number of participants and the value of the reward. It is necessary to evaluate the effect of this model on the number of participants and the total amounts of contributions. Thus, I experiment this model by computer simulation.

5.3.1 The Method of Experiments

The method of experiments is as follows.

I assume that the abilities are according to uniform distribution (0.5,1) and the number of users who might participate the contest are according to the value of the reward. I call these users “potential participants”. Hence, in case that there are p potential participants when $V = 1$, the number of potential participants in the first stage is $p(1 - \alpha)$ and the number of potential participants in the second stage is $p\alpha$. The potential participants in the first stage will participate if their contributions in the first stage are positive, by equation (42). The potential participants in the second stage will participate if their contributions in the second stage are positive, by equation (44). Thus, to determine whether user i participate the contest, order all participants by their abilities from lowest to highest. Then, user i would participate if and only if:

$$c_i < \frac{\sum_{k=1}^{i-1} c_k}{i-2} \tag{51}$$

By equation (39) and (41). $\sum_{k=1}^{i-1} c_k$ is the sum of values of abilities of which the participants are the lower than c_i (This means that the abilities of the participants at this point are higher than user i.). In case that all potential participants' ability are equal, all potential participants participated the contest. Otherwise, the low skilled potential participants would not participate the contest.

In situation where there are some new entrants, there is existing user i who exerted zero contribution ($x_{i2} = 0$) in the second stage. In this case, some utilities of existing user or new entrants would be negative by calculating from equation (28) and (29). Thus, these users are excluded from the set of potential participants, because they have no incentive to contribute. By repeating these processes, I specified the users who have positive contributions and utilities from the set of potential participants.

Under such conditions, I compared a case that no entrants exist (one-stage contest) with a case that new entrants existed and analyzed the number of the participants and the amount of contributions. Moreover, I analyzed the effect of α (distribution rate between first and second stage) and δ (carryover rate of contributions in the first stage) for the number of the participants and the amounts of contributions.

Algorithm 1 shows the process of determining the participants based on equation (51). Algorithm2 shows the process of calculating the amount of contribution in case that no entrant. V means the value of reward, *base_num* means the potential participants when $V = 1$ and *calculate_x* is a function to calculate the amount of each user's contribution by equation (25).

Algorithm 1 Specify participate users

```
procedure specifyParticipateUsers(participate_users, users, ability_sum)
    i ← 1
    if participate_users == nil then
        participate_users ← users[1], users[2]
        i ← 3
    end if
    while i < users.length do
        if users[i].ability < abilitySum/(i−2) then
            abilitySum ← abilitySum + users[i].ability
            participateUsers.append(users[i])
        end if
    end while
    return participate_users
end procedure
```

Algorithm 3 shows the process of calculating the amount of contribution in case that new entrants exist. *calculate_x1* is a function to calculate the amount

Algorithm 2 calculate optimal contributions with no entrants

Require: $V, base\_num$

{Initialization}

$users \leftarrow uniformSortedUserGenerator(base\_num * V)$

$participate\_users \leftarrow nill$

$ability\_sum \leftarrow 0$

{specify participate users}

$participate\_user \leftarrow$ SPECIFYPARTICIPATEUSERS($participate\_user, uses, ability\_sum$)

{calculate contributions}

for $i = 0 \ldots participate\_users$.length **do**

  $participate\_users[i].calculate\_x(V, participate\_users$.length, $abilitySum)$

end for

return $participate\_users$

of each existing users' contribution based on equation (36). $calculate\_x2$ is a function to calculate the amount of each existing users' and new entrants' contributions based on equation (37) and (38) respectively. $checkUtility()$ is a function to calculate the utility of each user by the number of participants and the total amount of contributions at that point and to exclude potential users who have negative utilities. This function returns true if there are potential users who are excluded, else false.

The result of experiments differs from the set of potential participants because I assumed that the abilities of potential participants are according to a uniform distribution. Therefore, I run a simulation a thousand times and then calculate the average values.

5.3.2 Result of Experiments

I show the experimental results.

a)A case that no entrants exist

I examined the number of participants and the amount of contributions where $V = 1$ and $base\_num$ is set to $100, 1000, 10000$ in a case that no entrants exist. I assume that the abilities of potential participants are according to uniform distribution.

Algorithm 3 calculate optimal contributions with new entrants.

```
Require: V, base_num, α, δ
  {Initialization}
  stage1users ← uniformSortedUserGenerator(base_num ∗ (1 − α)V)
  stage2users ← uniformSortedUserGenerator(base_num ∗ αV)
  participate_users ← nill
  ability_sum ← 0
  {specify participate users of first stage}
  participate_users
  ← specifyParticipateUsers(participate_user, stage1uses, ability_sum)
  while checkUtility(participate_user) do
     {calculate stage1 optimal contributions of 1st stage users}
     for i = 0 ... participate_users.length do
        participate_users[i].calculate_x1(V, participate_users.length, abilitySum, α, δ)
     end for
     {specity participate users of second stage}
     participate_users
     ← specifyParticipateUsers(participate_user, stage2uses, ability_sum, α, δ)
     {calculate stage2 optimal contributions of 2nd stage users}
     for i = 0 ... participate_users.length do
        participate_users[i].calculate_x2(V, participate_users.length, abilitySum)
     end for
  end while
  return participate_users
```

Figure 4a shows the effect of increasing the potential participants *base_num* in the case of no new entrant. The horizontal axis represents the number of potential participants *base_num* while the vertical axis represents the number of users. If *base_num* increases, the number of actual participants also increases but the increase is not linear to the increase of *base_num*.

Figure 4b shows the total amount of contributions. The horizontal axis represents the number of potential participants *base_num* while the vertical axis represents the total amount of contributions. If *base_num* increases, the total amount of contributions slightly increases but the difference remains small.

Figure 5 shows the relationship between abilities and the amount of contributions where *base_num* is set to 100,1000,10000. If *base_num* increase, the abilities of participants increases, but the amount of contribution by each participant decreases. This is because the case of 10000 is more competitive and the participation condition becomes harder. Thus, each users the contribution, compared to the case of $base\_num = 100$ or 1000 in order to satisfy the participation restriction by equation (51). For example, if $base\_num = 100$, potential participants who have abilities above 0.56 would participate the contest. While if $base\_num = 1000$, potential participants who have abilities above 0.52 would participate the contest. According to the above reasons, I could explain that the increase in the number of actual participants is not linear to the increase of *base_num*.

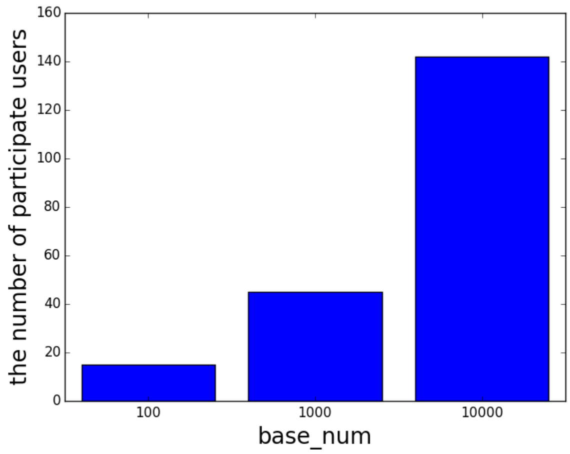
I summarize the result of the case of no new entrant as follows:

- If the number of potential participants increases, the number of actual participants also increases but the increase is not linear to the increase of potential participants.
- If the number of potential participants increases, the abilities of actual participants are higher.
- If the number of potential participants increases, the amount of contribution of each user decreases.
- If the number of potential participants increases, the total amount of contributions of all actual participants increases.

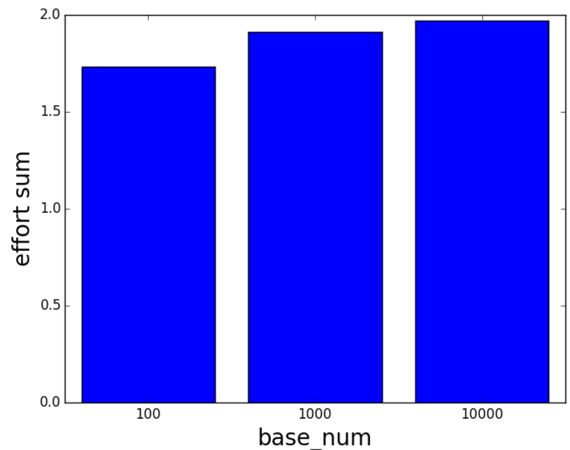
b)A case that new entrants exist

Next, I examined that the case that new entrants exist, where $V = 1$ and the number of potential participants *base_num* is set to either of $100, 1,000, 10,000$.

Figure 6 shows the total number of uses, where α changes over $[0, 1]$ and δ changes over $[0, 0.4]$. The horizontal axis represents the value of α while the vertical axis represents the total number of users, i.e., the number of users included in the second stage. Figure 6a, 6b, 6c correspond to the cases that *base_num* is set to $100, 1000, 10000$, respectively.



(a) The number of actual participants



(b) The total amount of contributions

Figure 4: The effect of increasing the potential participants *base_num* in the case of no new entrant

$\alpha = 0$ corresponds to the case that all the value is allocated the award in the first stage. If $\delta = 0$, there is no carryover of contributions to second stage. In this case, new entrants have no incentive to participate the second stage (their utility would be negative), thus, it is identical to the case of no new entrant. Compared with the case that no new entrants exist by Figure 4a, the number of users is almost the same value in the each *base_num*. If $\delta > 0$, there are some carryover of contributions to second stage. In this case, the

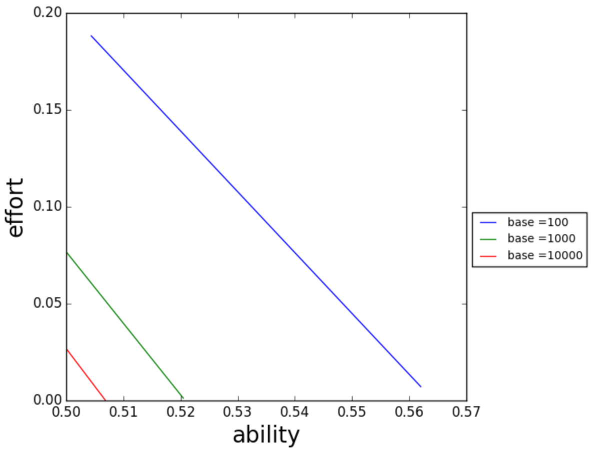


Figure 5: The relationship between abilities and the amount of contributions with different *base_num*

optimal contributions of existing users by equation (39) are higher than the case of $\delta = 0$. As the result, the number of users whose utilities are negative increases and the number of users who participate the contest decreases.

$\alpha = 1$ corresponds to the case that the all the value is allocated to the reward in the second stage. Therefore, the number of participants is nearly equal to the number in case that no new entrants.

For the case of $0 < \alpha < 1$, a reward is given to a user in each stage, and it is identical to the case of no new entrant. The following profiles can be implied from Figure 6

- The number of participants is the largest where α is slightly larger than 0.5.
- The number of participants decreases if δ increases.

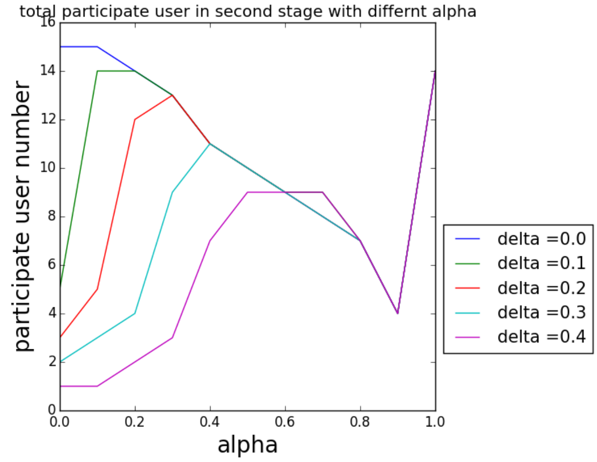
To examine why such profiles are obtained, I show the number of existing users(Figure 7a), and the number of new entrants(Figure 7b) in Figure 7 , which is the case of $base\_num = 10,000$. The sum of the existing users and the new entrants is the same as that shown in Figure 6c. First, I analyze the number of existing users. Except the case of $\delta = 0$, as α increases, the number of the incumbents first increase, and then decreases, and finally increases. Here, decreasing the value of the reward in the first stage means increasing the value of the reward in the second stage. However, the contest in the second stage is more competitive due to the new entrants. Thus, intuitively, it is natural to consider that the number of users decreases if the value of the reward in the first stage decrease. Why does it happen that the number of the existing users increases, as α increases from 0? This can be explained as follows. If δ increases, the equilibrium contribution in first stage increases to take advantage of carryover. For middle-ability user have to increase the amount of contributions, as high-ability users do. Otherwise, the middle-ability users become difficult to win the reward in the first stage. Investing more contributions in the first stage makes the equilibrium contribution in the second stage a negative value. However, the negative contributions are not allowed, and the expected utility of the middle-ability users become less than zero. Thus, these middle-ability users quit participating in all the contests. This explanation is supported by the fact that the number of existing users decreases if δ becomes larger.

For the new entrants, as α increases, the number of the new entrants first increases, and then decreases, and finally increases. Intuitively, it is natural to consider that the number of users increases if the value of the reward in second stage increases. Why does it happen that the number of new entrants decreases at $\alpha = 0.9$? This can be explained as follows. At $\alpha = 0.9$, the number of existing users decreases because the value of the reward in the first stage is small. However, such thinness of the competitors, as well as the large value of the award in the second stage, makes the users invest more contribution because they are quite high-ability users. As shown below, this can be confirmed from Figure 9. Thus, the number of the new entrants decreases at $\alpha = 0.9$. As mentioned above, $\alpha = 1$ corresponds to the case that the all the value is allocated to the award in the second stage. Thus, the number of participants is nearly equal to the number in case that no new entrants.

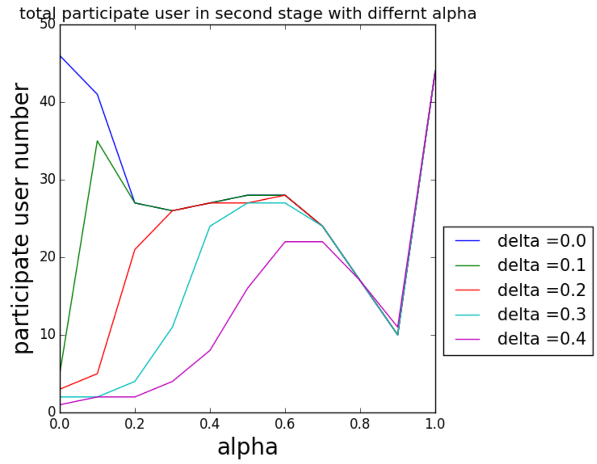
Next, I examine the total amount of contributions. Figures 8 shows the total amounts of contributions, where α changes over $[0, 1]$ and δ changes over $[0, 0.4]$. The horizontal axis represents the value of α while the vertical axis represents the total amount of contributions, i.e., the sum of the total contributions in the first stage and the total contributions in the second stage. Figure 8a, 8b, 8c correspond to the cases that $base\_num$ is set to $100, 1,000, 10,000$, respectively. From these graphs, I can learn that

- If δ increases, the total contributions decrease.
- If α increases, the total contributions decrease.
- if $base\_num$ increases, the effect of changing δ and α increases.

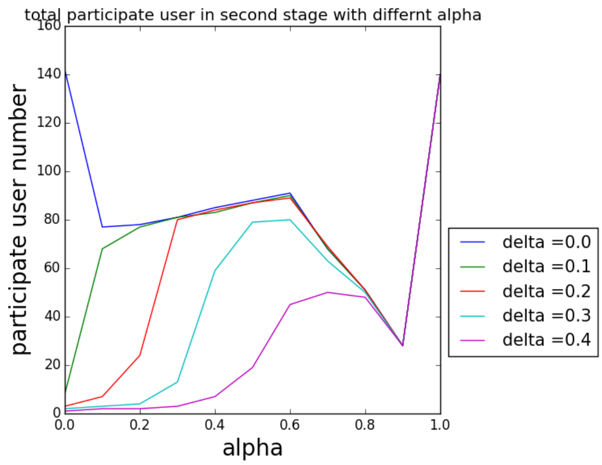
To examine the effect of changing δ and α, I show the total amount of contributions by the existing users in first stage(Figure 9a), that by the existing users in second stage(Figure 9b), and that by the new entrants in second stage(Figure 9c), which is the case of $base\_num = 10000$. Figure 9a shows that the larger δ yields the fewer contribution if α is small. This is because the larger equilibrium contribution causes more potential participants to suffer a loss by actual participation. Figure 9b shows that as the value of the reward in second increases, the total amount of contributions first increases, and then decreases. The reason of this decreasing is that thinness of competitors brings more competitiveness. Figure 9c shows that the larger δ reduce the total amount of contributions by the new entrants at the lowerα. This is natural because the larger δ makes the existing users more competitive.



(a) $base\_num = 100$

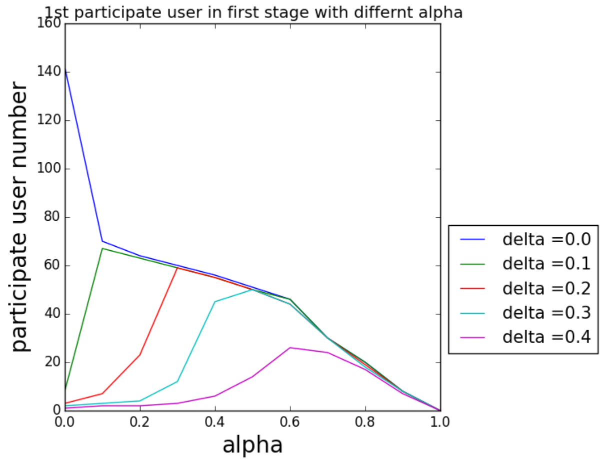


(b) $base\_num = 1000$

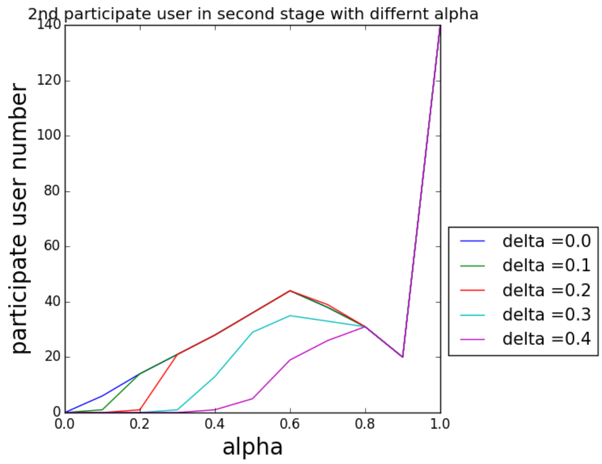


(c) $base\_num = 10000$

Figure 6: The number of participants with different α

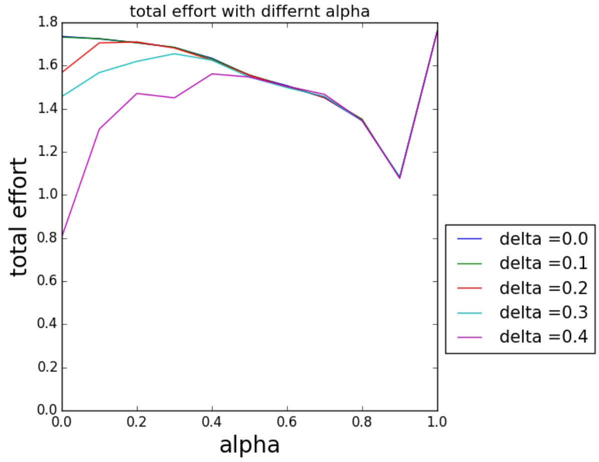


(a) The number of existing user in first stage

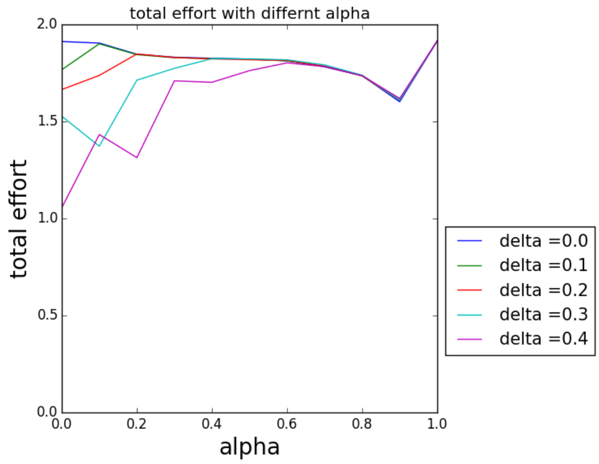


(b) The number of new entrants in second stage

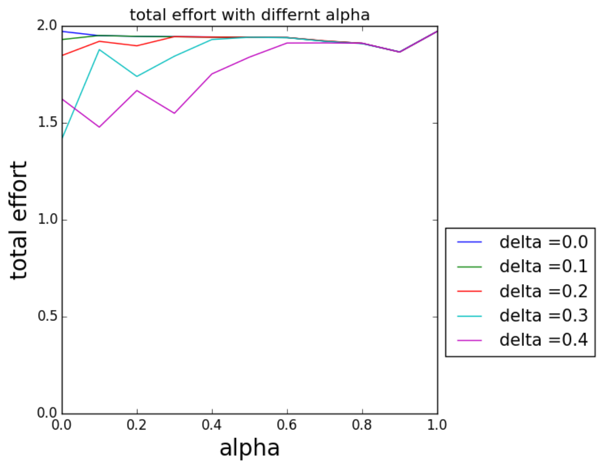
Figure 7: The number of participants in first and second stage in case that $base\_num = 10000$



(a) $base\_num = 100$

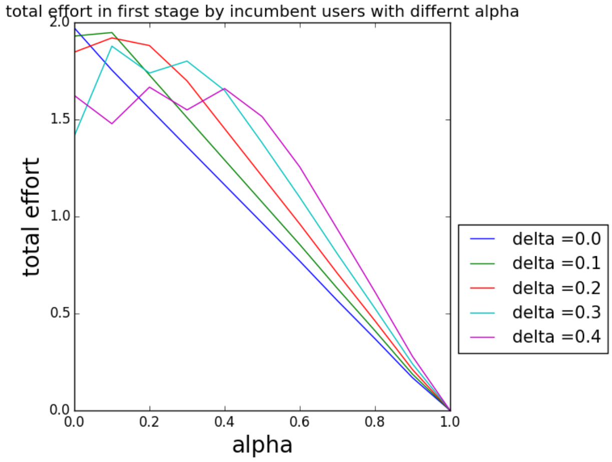


(b) $base\_num = 1000$

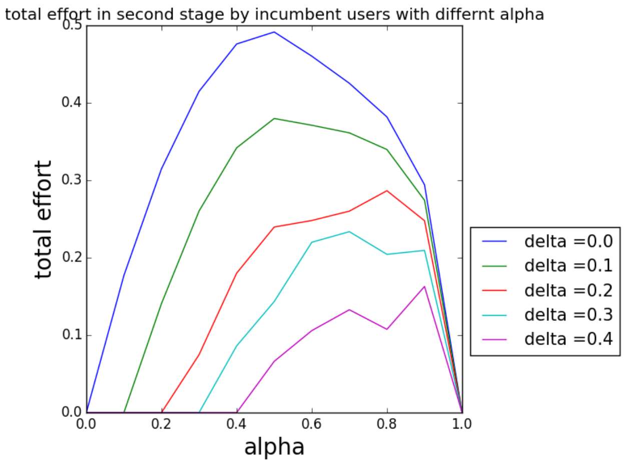


(c) $base\_num = 10000$

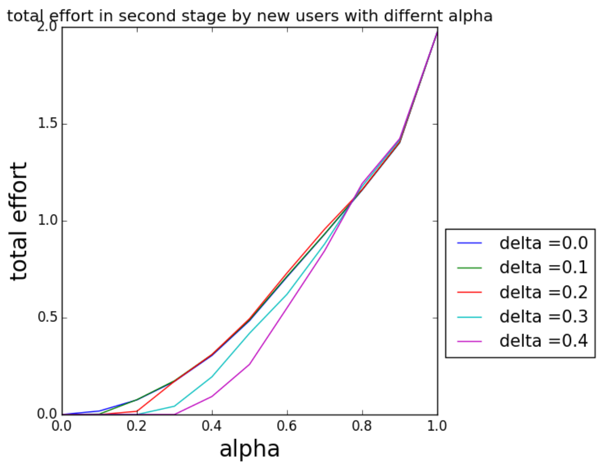
Figure 8: The total amount of contributions with different α



(a) Total amount of contributions of existing users in first stage



(b) Total amount of contributions of existing users in second stage



(c) Total amount of contributions of new entrants in second stage

Figure 9: The total amounts of contributions with different α in each stage in case that $base\_num = 10000$

Chapter 6 Discussion

In this chapter, I discuss the model based on order statics and the model based on rent-seeking theories.

In the model based on order statics, assumed that the distribution of users' abilities and the number of the users are common knowledge, the rewards are allocated in the order of the amount of the contribution and the designer could be set q rewards. While in the model based on rent-seeking theories, assumed that the abilities of the users and the number of the participants are common knowledge, a reward is stochastically allocated according to the amount of contributions and the designer could be set only one reward.

Table 8 shows the difference by these two models. Thus, it is difficult to evaluate by the same indicator because these two models are based on the different assumptions.

The model based on order statistics deals with a special situation where there is one existing user. Thus, it is difficult to apply this model to the real Q&A communities' situation directly. To explain Q&A communities where are mixed existing users and new entrants, it is necessary to model a situation where there are some existing users.

Otherwise, the model based on rent-seeking theories deals with a situation

Table 8: A comparison of two models

Model	Method of allocating reward	The number of rewards	Common knowledge
The model based on order statics	The order of the amount of the contribution	q	The distribution of users' abilities, number of users
The model based on rent-seeking theories	Stochastic according to the amount of the contribution	1	Abilities of users, the number of participants

where there are mixed existing users and new entrants. The situation where a reward is stochastically allocated according to the amount of contribution corresponds to, e.g., the situation where users compete for "Best Answer". However, there may be few Q&A communities where the carryover rate of contributions (δ) and the distribution rate of reward for each stage (α) could be set directly.

Considering a situation where those who have high reputation points would be easily trusted, if their reputation points are conspicuously displayed in near their answers, it can be said that δ is set large. For example, in StackOverflow, users' reputation points and obtained badges are displayed in near their answers, so that it can be said that δ is set large. While, if their reputation points are not conspicuously displayed in near their answers, it can be said that δ is set small. For example, in Yahoo!Answer, users' reputation points are conspicuously not displayed in near their answers so that it can be said that δ is set small.

Considering a situation where the badges are corresponded to social status and the more the number of users in the community become, the larger the value of badges is, it can be said that $\alpha > 0.5$. StackOverflow and Yahoo! Answer introduce badge system so that they are in this situation. While, if the badges are allocated for the users who fulfill the specific criteria, new entrants could not get the first badge, it can be said that $\alpha < 0.5$.

In the model of rent-seeking theories, I regard Q&A a community as a long contest to compete for many badges. From the result of analyzing this model, I indicate following instructions of incentive design for Q&A communities to Q&A communities.

1. Better not to carry over reputation points (reputation points should be reset in a fixed term)
2. Better to set larger reward in later than earlier.

For 1, I could arrive at this conclusion based on the fact that if δ decreases, the number of users and the total amount of contributions increase. For 2, I could arrive at this conclusion based on the fact that the number of participants is the largest where α is slightly larger than the half. As above mentioned, StackOverflow can be regarded as the situation where α and δ is large. In this situation, it may be difficult for new entrants to have the incentive to participate the community because δ is set larger. Yahoo! Answers can be regarded as the situation where α is large and δ is small. In this situation, it may be easy for new entrants to have the incentive to participate.

In this study, I could indicate one of the design guidelines for Q&A communities. However, it is a problem that I could not associate these situations with real data.

Chapter 7 Conclusions

In this study, I aim to design incentive suitable for Q&A communities where there are unspecified new entrants. Some of these Q&A communities relying on users' contributions employ some form of virtual rewards to increase engagement and to motivate their users to participate actively in the communities.

A Q&A community, which employs virtual rewards, could be regarded as a contest to compete for these rewards. However, it is difficult to model a Q&A community within the framework of existing contest theories, because new entrants are not considered in existing contest theories.

To solve above problems, I propose two models for Q&A communities by extending existing contest models. The first model is a model based on order statics. In this model, the rewards are allocated in the order of the amount of the contributions. I formulate the optimal contribution function to maximize each user's expected utility, assumed that the number of users, the distribution of abilities, and the number and the value of rewards are known. The second model is a model based on rent-seeking theories. In this model, the reward is stochastically allocated according to the amount of the contributions. Just like the first model, I formulate the optimal contribution function to maximize each user's expected utility, assumed that the number of users, the abilities of all users and the value of the reward. In this model, there is only one reward.

Next, I consider a Q&A community to compete for the prizes as a two-stage contest and extend above two models assumed that there are some users who participate the contest from the second stage.

For the model based on order statics, to simplify, I model the situation where there is one existing user from the first stage, there are some new entrants from the second stage, and the contributions of the existing user are carried over. I assume that the number of users and the distribution of abilities are common knowledge just like the case of no entrant. In this situation, I formulate the optimal strategies of existing and new enter users respectively to maximize their expected utilities.

For the model based on rent-seeking theories, I model the situation where there are existing users from the first stage, there are some new entrants from the second stage, and a part of the contributions of the existing user are carried over. I assume that the number of existing and new enter user, the abilities of all users and the value of the reward. In this situation, I formulate the optimal strategies of existing and new enter users respectively to maximize their expected utilities.

Next, to evaluate the effect of the incentive design on users' behavior in the case that new entrants exist on two proposed models, I experiment these models by computer simulation. About the model based on order statics, I confirm that the number of rewards and the ability of the existing user affect the total amount of contributions. About the model based on rent-seeking theories, I confirm that the rate of carryover from the first stage to the second stage affect the number of participants and the total amount of contributions

Finally, I analyze these results of experiments and indicate instructions of incentive design for Q&A communities. To activate communities, I propose following design guidelines.

1. Better not to carry over reputation points (reputation points should be reset in a fixed time)
2. Better to set larger reward in later than earlier.

The contributions of this study are followings:

Providing the model for Q&A communities considering new entrants

I construct two models for Q&A communities by extending the existing contest theories. I regard a Q&A community as a two-stage contest, and I formulate the optimal contribution functions for existing users and new entrants

Clarifying the effect of incentive design on Q&A communities

To analyze the effect of incentive design on the two-stage contests, I conduct some experiments based on simulation. As the result, I confirm that if the carryover rate of the contributions decreases, the number of users and the total amount of the contributions increase. Also, I indicate the design guidelines for soliciting contributions to Q&A communities.

In this study, these models are simulated on the computer, but I do not associate these models with real data. To apply these models to the real situation, I need to associate with real data and to evaluate in real environments.

Acknowledgments

I would like to express his gratitude to the supervisor, Associate Professor Shigeo Matsubara at Kyoto University, for his consecutive instruction,pertinent advice and earnest argument. I would like to express sincere appreciations to the advisers, Professor Hirokazu Tatano at Kyoto University, Associate Professor Hiroshi Ueda at Kyoto University for his valuable advice. I would like to tender his acknowledgment to Professor Toru Ishida at Kyoto University. Finally, I would like to thank all member of Ishida and Matsubara laboratory for their various support and discussion.

References

[1] Harper, F. M., Raban, D., Rafaeli, S. and Konstan, J. A.: Predictors of answer quality in online Q&A sites, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 865–874 (2008).

[2] Pomerantz, J., Nicholson, S., Belanger, Y. and Lankes, R. D.: The current state of digital reference: Validation of a general digital reference model through a survey of digital reference services, *Information Processing & Management*, Vol. 40, No. 2, pp. 347–363 (2004).

[3] Shah, C., Oh, S. and Oh, J. S.: Research agenda for social Q&A, *Library & Information Science Research*, Vol. 31, No. 4, pp. 205–209 (2009).

[4] Vargo, A. and Matsubara, S.: Are Good Questions Scarce in Domain Specific Q&A, 情報処理学会第 77 回全国大会, Vol. 6, p. 07 (2015).

[5] De Alfaro, L., Kulshreshtha, A., Pye, I. and Adler, B. T.: Reputation systems for open collaboration, *Communications of the ACM*, Vol. 54, No. 8, pp. 81–87 (2011).

[6] Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G. and Hartmann, B.: Design lessons from the fastest q&a site in the west, *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, pp. 2857–2866 (2011).

[7] Correa, D. and Sureka, A.: Chaff from the wheat: Characterization and modeling of deleted questions on stack overflow, *Proceedings of the 23rd international conference on World wide web*, ACM, pp. 631–642 (2014).

[8] Dechenaux, E., Kovenock, D. and Sheremeta, R. M.: A survey of experimental research on contests, all-pay auctions and tournaments, *Experimental Economics*, Vol. 18, No. 4, pp. 609–669 (2015).

[9] Baye, M. R., Kovenock, D. and De Vries, C. G.: The all-pay auction with complete information, *Economic Theory*, Vol. 8, No. 2, pp. 291–305 (1996).

[10] Moldovanu, B. and Sela, A.: The optimal allocation of prizes in contests, *American Economic Review*, pp. 542–558 (2001).

[11] Tullock, G.: The welfare costs of tariffs, monopolies, and theft, *Economic Inquiry*, Vol. 5, No. 3, pp. 224–232 (1967).

[12] Tullock, G.: Efficient rent seeking, *Efficient Rent-Seeking*, Springer, pp. 3–16 (2001).

[13] Anderson, L. R. and Stafford, S. L.: An experimental analysis of rent seeking under varying competitive conditions, *Public Choice*, Vol. 115, No. 1-2, pp. 199–216 (2003).

[14] Schmitt, P., Shupp, R., Swope, K. and Cadigan, J.: Multi-period rent-seeking contests with carryover: Theory and experimental evidence, *Economics of Governance*, Vol. 5, No. 3, pp. 187–211 (2004).

[15] Easley, D. and Ghosh, A.: Incentives, gamification, and game theory: an economic approach to badge design, *Proceedings of the fourteenth ACM conference on Electronic commerce*, ACM, pp. 359–376 (2013).

[16] Immorlica, N., Stoddard, G. and Syrgkanis, V.: Social Status and the Design of Optimal Badges.

[17] Balakrishnan, N. and Rao, C. R.: *Order statistics: theory & methods*, Elsevier Amsterdam (1998).

[18] Anderson, A., Huttenlocher, D., Kleinberg, J. and Leskovec, J.: Discovering value from community activity on focused question answering sites: a case study of stack overflow, *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 850–858 (2012).