Agent-Based Model of Q&A Community for Effective Pecuniary Payback System

Yuki Ogawa[1],Hitoshi Yamamoto[2],Isamu Okada[3],Hirohiko Suwa[1],Toshizumi Ohta[1]

[1] University of Electro-Communications, Graduate School of Information Systems
Chofu-shi, 81, 182-8585 Tokyo, Japan. ogawa@ohta.is.uec.ac.jp
[2] Rissho University, Faculty of Business Administration
[3] Soka University, Faculty of Business Administration

Abstract. Q&A communities have been developed. However, it is difficult to consider a design of the management systems effective for the community. We present an agent-based model, focusing on pecuniary payback, of a Q&A community, and we clarify the effective payback system policy for various community environments through simulation. Our results showed the following. In a system where pecuniary payback is low, there are many answers. In addition, in a system where pecuniary payback is high, quality of answers is high, and even if a questioner's question-effort is low (even if a questioner posts an ambiguous question), the community for which answers gather is formed.

Keywords: agent simulation, Q&A community, pecuniary payback, motivation

1 Introduction

1.1. Payback System in Q&A Community

Several knowledge-sharing communities such as Q&A communities have been developed (e.g., “Yahoo!Answers”[1], “Yahoo!Chiebukuro”[2] and “Hatena::Question”[3]). Q&A communities are places where users ask questions and others answer them. Yahoo Research has claimed that “Yahoo!Answers” is the next generation of search. It is a kind of collective intelligence - a searchable database of everything everyone knows. For continuous development of these Q&A communities, it is important to design an effective management system. In particular, a payback system that involves an answerer's motivation is important.

In an actual Q&A community, there are various payback system policies. For example, “Yahoo!Chiebukuro” uses a non-pecuniary payback system (e.g., a word of gratitude or marking of an answer as a “Best Answer”). On the other hand, “Hatena::Question” uses a pecuniary payback system called "Hatena Points". A non-pecuniary payback system has the advantage of lowering the payback cost to the answerer for a questioner, while a pecuniary payback system has the advantage of raising an answerer's motivation. However, the long-term effect of the difference in the amount of pecuniary payback on a community is not clear.

---

[1] Yahoo!Answers: *http://answers.yahoo.com/*
[2] Yahoo!Chiebukuro: *http://chiebukuro.yahoo.co.jp/*
[3] Hatena::Question: *http://q.hatena.ne.jp/*

1.2. Related Works

Studies have been conducted on a participant's motivation and on an answerer's answer category as part of the research on Q&A communities. However, in these studies, neither the interaction between a questioner and an answerer nor the effect that a payback system has on a community is clarified.

Miura [6] investigated the motives of participants in the “Yahoo!Chiebukuro”(one of the most popular Q&A community in Japan) by applying a questionnaire. As a result, Miura et al. classified the motives of participating answerers into four categories: (1) support motive, (2) reciprocity motive, (3) social motive, and (4) reward motive. This result shows that change of the motivation due to interaction between a questioner and an answerer must be considered in order to consider development of a Q&A community.

A recent study modeled the community participants’ action patterns as agents and modeled the development process of the community by agent simulation. Yamada [8] proposed a model that comprehensively depicted the open-source software development community and the Q&A community. However, there are various success indexes for a community [5], and to understand the growth mechanism of a community, detailed modeling of the community extracted for analysis is required.

In this work, to show clearly what kind of effect the payback system in a Q&A community has on that community, we present an agent-based model of a Q&A community that focuses on the payback system. Then, we clarify the payback system policy that is effective for each community environment through simulation. This study would be helpful for the administrator of a Q&A community to decide on a suitable payback system.

2 Agent-Based Model of Q&A Community

We present abstract modeling [4] that focuses on the payback system. Specifically, we focus on a Q&A community and present an agent-based model focusing on the mental motive that Miura et al. reported [6].

2.1. Knowledge Transactions between Questioner and Answerer

We model the interaction between a questioner and an answerer as three processes.

Question Post from Questioner. A questioner posts a question based on his or her knowledge. Because this question is expressed as text, the ambiguity of a question depends on the efforts of the questioner. Thus, we model a questioner as one who posts a question generated through questioner’s efforts to a community.

Answer Post from Answerer. An answerer posts an answer to a question that matches his or her own knowledge. The answerer puts effort into providing comprehensive information in text-form in answer to an ambiguous question. Thus, we model an answerer as one who chooses a question in consideration of answer’s knowledge property.

Payback Post from Questioner to Answerer. A questioner gives payback to the answerer. However, the answer may not necessarily be one needed by the questioner. Thus, we model a questioner as one who pays back an answerer who provides a matching answer.

2.2. User Model

We model a questioner and an answerer based on the knowledge transaction process described in Section 2.1. In addition, we refer to Axelrod's tag model [2], and model knowledge as that which expresses the personal knowledge property abstractly. Specifically, we express the knowledge property (e.g., such as knowledge about a genre) that an individual has by the bit string of 0 and 1.

Questioner's Knowledge Property (*Prop*^{Q}). Formula (1) shows questioner *i*'s knowledge property ($Prop^Q_i$). Each bit expresses questioner *i*'s knowledge property over one knowledge unit (e.g., such as knowledge about a genre).

$$Prop^Q_i = [x_1, x_2, \ldots, x_k] = [x_n \in \{0, 1\}] \quad \textbf{(1)}$$

Questioner's Question-Effort (*Effort*^{Q}). $Effort^Q_i$ shows the effort that questioner *i* put into writing a question, i.e., question-effort ($0 \leq Effort^Q_i \leq 1$). If questioner *i* has a low $Effort^Q_i$ value, questioner *i* generates unambiguous text (question $Info^Q_i$).

Questioner's Question (*Info*^{Q}). Formula (2) shows the question ($Info^Q_i$) that questioner *i* posts. We express $Info^Q_i$ as ambiguous question that is caused by the question-effort ($Effort^Q_i$) in correspondence to knowledge property ($Prop^Q_i$). The "*" in formula (2) means the ambiguity of a question, and each bit string of $Prop^Q_i$ changes to "*" by the probability of ($1\text{-}Effort^Q_i$). Specifically, formula (2) expresses that "*" is part of the state of being supplied with an answerer's efforts. For example, an answerer may give a comprehensive answer to an ambiguous question about PCs.

$$Info^Q_i = [x_1, x_2, ..., x_k] = [x_n \in \{0, 1, *\}] \quad \textbf{(2)}$$

Questioner's Payback Threshold (*Th*^{Q}). Th^Q_i shows questioner *i*'s payback threshold ($0 \leq Th^Q_i \leq 1$). Questioner *i* gives payback to an answerer when the value of a match between questioner *i*'s knowledge property ($Prop^Q_i$) and an answer ($Info^A$) exceeds questioner *i*'s payback threshold. In addition, the value of the match normalizes the Hamming Distance of questioner *i*'s $Prop^Q_i$ and an answer ($Info^A_j$) by Answerer *j* to a value between 0 to 1.

Answerer's Knowledge Property (*Prop*^{A}). Formula (3) shows answerer *i*'s knowledge property ($Prop^A_i$).

$$Prop^A_i = [x_1, x_2, ..., x_k] = [x_n \in \{0, 1\}] \quad \textbf{(3)}$$

Answerer's Answer-Effort (*Effort*^{A}). The effort put in by an answerer in producing an answer to a questioner's ambiguous question is expressed as answer-effort $Effort^A_i$ ($0 \leq Effort^A_i \leq 1$). If answerer *i* tries hard to produce an answer, the value of $Effort^A_i$ is high, and text (answer $Info^A_i$) is generated.

Answerer's Answer Threshold (*Th*^{A}). An answerer will answer a question that matches answer *i*'s knowledge. Answerer *i* posts an answer to a question when the value of a match between answerer *i*'s knowledge property ($Prop^A_i$) and a question ($Prop^Q_j$) exceeds answerer *i*'s answer threshold ($0 \leq Th^A_i \leq 1$). The value of the match normalizes the Hamming Distance of answerer *i*'s $Prop^A_i$ and a questioner *j*'s $Info^Q_j$ to a value between 0 to 1. In addition, because the "*" in the Th^A for $Info^Q_j$ means the ambiguity of a question, the Hamming Distance of "*" is calculated as 1.

Answerer's Answer (*Info*^{A}). Formula (4) shows the answer ($Info^A_i$) by answerer *i*. An answerer posts a comprehensive $Info^A_i$ to an ambiguous question corresponding to answer *i*'s answer-effort ($Effort^A_i$). Specifically, Answerer *i* generates $Info^A_i$ having a value with which only the rate of $Effort^A{}_i$ brought own knowledge $Prop^A{}_i$ close to each bit of a questioner's knowledge property $Prop^Q{}_j$ to the "*" in a question $Info^Q{}_j$. As in the previous example, an answerer tries hard to answer an ambiguous question about PCs and produces a comprehensive answer.

$$Info^A{}_i = [x_1, x_2, ..., x_k] = [x_n \in \{0, 1\}] \qquad \textbf{(4)}$$

Answerer's Support Utility (*AS*). We express the utility (AS_i) that Answerer *i* obtains by posting an answer ($0 \leq AS_i \leq 1$), based on the support motive [6].

2.3. Modeling of Pecuniary Payback System (*PP*)

Fehr [3] has classified utility into pecuniary utility and non-pecuniary utility. From this viewpoint, a user's utility in a Q&A community is a product of pecuniary payback and non-pecuniary payback. In addition, pecuniary payback is a variable that a community administrator can control. Thus, we model the pecuniary payback system in a Q&A community as *PP* ($0 \leq PP \leq 1$). If a questioner gives high payback and, specifically, an answerer receives high payback, the value of *PP* is high.

2.4. User's Utility Model

Questioner's Utility (U^Q). We express Questioner *i*'s utility $U^Q{}_i$ as a value in which question cost $Effort^Q{}_i$ and *PP* are subtracted from the degree of the match between knowledge property $Prop^Q{}_i$ and answer $Info^A{}_j$.

$$U^Q{}_i = match(Prop^Q{}_i, Info^A{}_j) - Effort^Q{}_i - PP \qquad \textbf{(5)}$$

Answerer's Utility (U^A). We express answerer *j*'s utility $U^A{}_j$ as a value in which the answer creation cost $Effort^A{}_i$ is subtracted from *PP*, and AS_j is added.

$$U^A{}_j = PP - Effort^A{}_j + AS_j \qquad \textbf{(6)}$$

2.5. Community Environment (similarity of the knowledge tendencies (*S*))

Currently, Q&A is used in various community environments, such as communities with many participants with various knowledge (e.g., “Yahoo!Chiebukuro”), communities with many participants with a similar knowledge tendency (e.g., a UNIX community), and communities with many participants with different knowledge tendencies (e.g., a knowledge management community).

We express a community environment as the similarity of the knowledge tendency between a questioner and an answerer. Specifically, we express *S* ($0 \leq S \leq 1$) as the expected value of the coincidence rate of the value of a questioner's knowledge property ($Prop^Q$) and an answerer's knowledge property ($Prop_A$) (formula (7)). In this formula, random variable *X* expresses the probability that each bit of $Prop^Q$ will take 1, and random variable *Y* expresses the probability that each bit of $Prop^A$ will take 1. *X* and *Y* are independent random variables. Furthermore, E_X is the expected value of *X*, and E_Y is the expected value of *Y*.

$$S = (1 - E_X)(1 - E_Y) + E_X E_Y \qquad \textbf{(7)}$$

2.6. User Breakaway and Entry

In an actual Q&A community, breakaway of participants with low motivation of to answer or ask and entry of new participants occurs. Therefore, in this model, a user with low utility breaks away and a new user enters. Specifically, low total utility users are replaced by a new user.

3. Experiments and Results

3.1. Outline of Experiments

We conducted a simulation experiment. The conditions of the simulation are shown below. The simulation repeated the following Steps 1–3 500 times.

Step 1: Action Selection

- A questioner posts a question to a community.
- An answerer calculates the degree of match between answer's knowledge and a question and then answers questions for which the degree is answer's threshold.
- A questioner calculates the degree of match between questioner's knowledge and an answer and gives payback for answers that are over questioner's threshold.

Step 2: Calculation of Utility

- The utility of a questioner and an answerer is calculated (formulas (5) and (6)).

Step 3: Breakaway and Entry

- A low-rank user (10% of participants with the lowest total utility in the latest 5 terms) with low total utility is replaced by a new user.

Table 1 shows the parameters used for the experiment. The results are the average of 100 simulations in which the seed (a random number) was changed.

3.2 Comparison of Proposed Model and Actual Q&A

To investigate the validity of our proposed model, we compared the participant action of the proposed model with the participant action in actual Q&A communities ("Yahoo!Chiebukuro" and "Hatena"). The plots in Fig. 1 show one answerer's answer-effort ($Effort^{A}$) and the value of the acquisition rate of payback. In addition, we used the data of "Yahoo!Chiebukuro," which was provided by the National Institute of Informatics, and the crawling data of "Hatena". Moreover, the answer-effort in an actual Q&A community is assumed on the basis of the number of letters.

In Fig. 1, the user action in our model and the user action in an actual Q&A community correspond. This result indicates that our proposed model can depict an actual Q&A community. Section 4.1 will examine again the mechanism that brought about this correspondence.

- Data term: "Yahoo!Chiebukuro" (2004/10–2005/10), "Hatena" (2009/12–2008/12)
- Answerers: 1,000 (randomly chosen)

Table 1. Parameter of Simulation

Variable		Value
Environme	Simulation times	500 times
nt	Agents	Questioners 100, Answers 100
Questioner	Questioner *i*'s Knowledge Property($Prop^{Q}_{i}$)	uniform random numbers [0,1]
	Questioner *i*'s Question Effort ($Effort^{Q}_{i}$)	uniform random numbers [0,1]
	Questioner *i*'s Payback Threshold (Th^{Q}_{i})	uniform random numbers [0,1]
Answer	Answerer *i*'s Knowledge Property ($Prop^{A}_{i}$)	uniform random numbers [0,1]
	Answerer *i*'s Answer Effort ($Effort^{A}_{i}$)	uniform random numbers [0,1]
	Answerer *i*'s Answer Threshold (Th^{A}_{i})	uniform random numbers [0,1]
	Answerer *i*'s Support Utility (*AS*)	uniform random numbers [0,1]
Communit	Pecuniary Payback System (*PP*)	[0, 0.1, 0.2, …, 1.0]
y	Similarity of the Knowledge Tendencies(*S*)	[0.3, 0.5, 0.7]

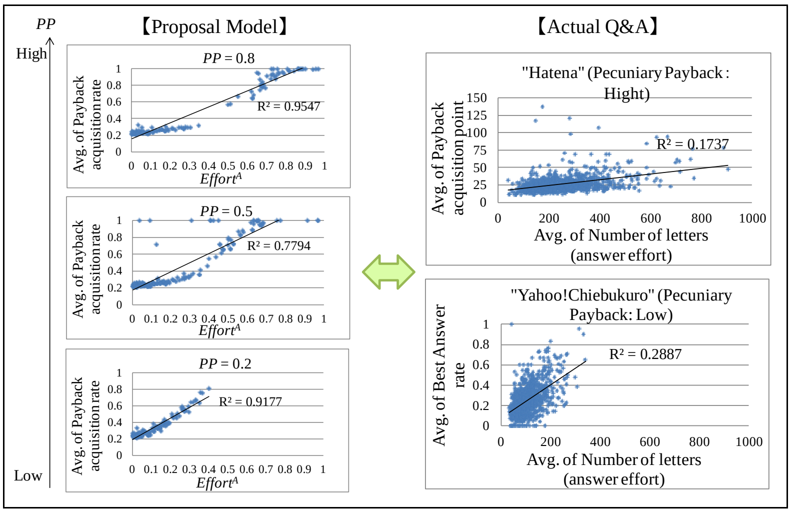

Fig. 1. Comparison of Proposed Model and Actual Q&A

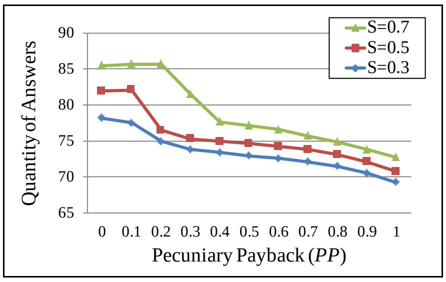

Fig. 2. Quantity of Answers

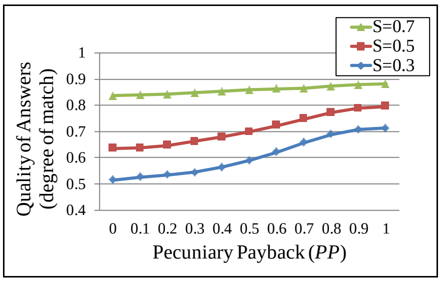

Fig. 3. Quality of Answers

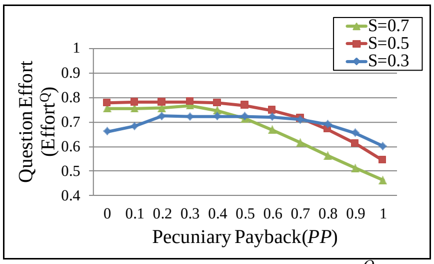

Fig. 4. Question-Effort ($Effort^Q$)

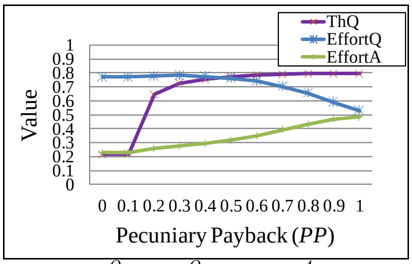

Fig. 5. Th^Q, $Effort^Q$ and $Effort^A$ in PP=0.5

3.3 Effect of Pecuniary Payback System

Quantity of answers. Figure 2 shows the quantity of answers (average of the quantity of answers to one question) for changing *PP*. The results show that a community with a high quantity of answers was achieved by the system in which pecuniary payback was low.

Quality of answers. Figure 3 shows the quality (average of the degree of match between a questioner's knowledge property and an answer) of the answers for changing *PP*. The results show that a community with high-quality answers was achieved by the system in which pecuniary payback was high.

Question-effort (ease of asking questions). Figure 4 shows the question-effort for changing *PP*. The results show that a community in which users tend to ask questions was achieved by the system in which *PP* was high.

Figure 5 shows the questioner's payback threshold and question-effort, and the answerer's answer-effort for changing pecuniary payback. The results show that the payback threshold and the answer-effort increased in proportion to the amount of pecuniary payback and that a questioner's question-effort declines in inverse proportion to the increase in an answerer's answer-effort. Furthermore, in an environment having participant similarity $S = 0.3$, a community in which users tend to ask questions was also achieved under a system where pecuniary payback was low.

4 Discussion

4.1. Quantity and Quality of Answers

The results in Figs. 2 and 3 show the following. In a community with low pecuniary payback, although there is a large quantity of answers, the quality of those answers is low. In contrast, in a community with high pecuniary payback, although the quality of answers is high, there are few of them.

The results in Fig. 5 show the mechanism by which pecuniary payback causes the quantity and quality of answers to change. In our model, pecuniary payback is the cost for a questioner. Therefore, in the community system where pecuniary payback is high, a questioner whose payback threshold is higher than a questioner with a low payback threshold will obtain a high profit. Then, a questioner with a high payback threshold will continue in the community. As for answerers, because an answerer with a high answer-effort value receives payback from a questioner, this answerer continues in the community. For this reason, in a community with high pecuniary payback, the quality of answers is high because the effort of answerers is high. However, the quantity of answers decreases because the value of a questioner's payback threshold becomes high (i.e., it becomes difficult for an answerer to receive payback).

4.2. Question-Effort (Ease of Asking Questions)

The results in Fig. 4 show that a community in which questions tend to be asked was achieved by the system in which pecuniary payback is high. In this case, an answerer's efforts become high due to the mechanism described in Section 4.1, and answers can be gathered even if a questioner does not have a high question-effort value. This result supports the effect which the administrator in an actual Q&A community expects [7].

Moreover, the results in Fig. 5 show that not only a questioner with a high question-effort value but also a questioner with a low question-effort value received answers in the community having similarity S=0.3. In such a community, contact probability with an answerer with whom a particular knowledge tendency is similar is low. Therefore, a questioner's profit acquisition from obtaining a match with low probability, rather than paying the cost of higher question-effort and obtaining an answer, is strong. For this reason, a questioner with a low question-effort value is considered to have continued in the community. In addition, in a community system where pecuniary payback is low, because the profit acquisition from matching is big, the above situation is thought to be reflected strongly.

4.3. Payback System Policy in Q&A Community

In a community with high similarity of knowledge tendency between a questioner and an answerer, if an administrator considers the quantity of answers important, the policy for an effective payback system would be to set the pecuniary payback low. However, if an administrator considers as important the quality of answers and the ease of asking questions, the policy for an effective payback system would be to set the pecuniary payback high. In contrast, in a community with low similarity of knowledge tendency between a questioner and an answerer, because it is easy to ask a question even if pecuniary payback is low, the policy for an effective payback system would be to set the pecuniary payback low.

5 Conclusion and Future Work

In this study, to show clearly what kind of effect the payback system in a Q&A community has on the community, we constructed an agent-based model of a Q&A community that focused on the payback system. Then, we clarified the payback system policy that is effective for each community environment through simulation.

The limitation of this study is that our model focused on pecuniary payback. In an actual Q&A community, an answerer's motivation may not only be pecuniary payback but also reputation etc. Our future task is to consider modeling designs of management systems other than pecuniary payback.

References

1. Adamic, L. A, Zhang, J. and Ackerman, M. S.: Knowledge sharing and Yahoo Answers: Everyone knows something. *Proc. 17th Int. Conf. on World Wide Web* (*WWW2008*) (2008)
2. Axelrod, R.: The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution* 41 (1997) 203-226
3. Fehr, E. and Schmidt, K. M.: A Theory of Fairness, Competition and Cooperation. *The Quarterly Journal of Economics* (1999) 817-868
4. Gilbert, G. N. and Troitzsch, K. G.: Simulation for the Social Scientist. Open University Press (1999)
5. Lin, H., Fan, W., Wallace, L. and Zhang, Z.: An empirical study of web-based knowledge community success. *Proc. 40th Int. Conf. System Sciences* (*HICSS'07*) (2007) 178c
6. Miura, A. and Kawaura, Y.: Why do people join Web-based knowledge-sharing communities? : Analysis on questioning and answering behaviors. *Japanese Journal of Social Psychology* 23(3) (2008) 233-245
7. Orita, A.: Why do people teach each other? (2). IT-PLUS (2009) http://it.nikkei.co.jp/internet/news/index.aspx?n=MMIT2E000003022009
8. Yamada, K., Nakakoji, K. and Yamamoto, Y.: Modeling of Incentive Mechanisms for an Online Community, *Proc. Joint Agent Workshops and Symposium* (*JAWS2008*) (2008)