A Design for a Document Evaluation System Allowing Critical Evaluation with Ordered Discussion

Atsushi Iwai, Hiromu Masaki and Yosuke Yoshizaki

***Abstract*—Recently many communication systems like SNS do not completely guarantee a user's anonymity, and this feature of their design helps to maintain the orderliness of user interactions. However, from the viewpoint of decision making or judgment, not guaranteeing anonymity gives rise to a problem concerning the value of any critical comments, as some users may refrain from making such critical comments for fear of provoking the negative reactions of others. This paper proposes a new design for a document evaluation system that encourages users to freely submit critical comments so as to increase the information content and value of any communication, and that also maintains the orderliness of interactions.**

***Index Terms*—Critical Evaluation, Ordered Discussion, Document Selection, Anonymity, Communication System.**

I. Introduction

SNS is a communication system that provides limited user anonymity or privacy and has recently become widespread in Japan. This characteristic of limited anonymity is in contrast to the current complete anonymity found in conventional BBS where any user can freely write a comment, secure in the knowledge that their privacy is guaranteed; unfortunately this often leads to context-less, meaningless and disorderly disputes and is a headache for system managers, resulting in the increased cost of system management. Inclusion of limited anonymity in the design of SNS is widely recognized to be contributing to the maintenance of well-ordered discussion and many schools, companies and local governments are willing to employ small scaled SNS in their organizations as an efficient type of communication tool to keep management costs down.

However, one drawback of limited anonymity is that it can also make users refrain from sending any critical messages that may cause a negative reaction from other users. For example, a SNS user may refrain to input a comment in the system to express their critical view of a movie, especially when the movie is receiving a highly favorable reception in the real world. The user may need some "courage" to openly express their views as their comment may cause a disturbance or negative reactions from other "close" users in the system.

High level decision making or judgment often needs frank communication involving critical evaluation. From this point of view, it would be reasonable to create a communication system that doesn't restrain, but encourages the free expression of critical evaluation; however, it is also desirable that the system has a mechanism that helps to prevent chaotic situations full of context-less and slanderous statements that hinder cool-headed and open discussion.

This paper presents a new design for a system that can support users to make critical evaluations and comments and at the same time maintains autonomous order among participants; particular attention is paid to usage in fields where communication involving open and critical evaluation is essential, such as book-reviewing. We will illustrate typical aspects of the problem and the design to solve them using a specific example of document evaluation where both each individual and a whole group have to make decisions concerning document selection.

We selected document evaluation as our subject, as this would be a less emotive subject for participants to discuss than other more volatile subjects such as events in the political and financial fields. That is to say, we evaluated the subject of document evaluation as an adequate first subject, presenting a suitable degree of difficulty to establish a method for the maintenance of well-ordered discussion.

As suitable candidates to benefit from the method, this study focuses on small organizations whose population ranges from tens to hundreds. They are expected to be organizations that have clear borders to separate members who belong to the organizations and outsiders, such as corporations or schools in the real world.

To support critical evaluation, the system guarantees anonymity to avoid the SNS problem mentioned above, and also employs another mechanism for users to allow them the freedom to make and be more involved in the outputting of critical comments. On the other hand, to maintain orderly discussion, an approach employing the "establishment of group rules and the automation of their enforcement" is used. Although this paper illustrates the new system design by using a specific example of document evaluation, the design is introduced here simply as a prototype and is also concerned with its application in various other fields that need communication involving critical evaluation for high level decision making.

The structure of this paper is as follows: the second section illustrates an example of document evaluation and sets out the typical problems that need to be solved in the system

Manuscript received December 8, 2008. Part of this study was supported by grants from the Japan Society for the Promotion of Science, KAKENHI18700246.

A. Iwai is with the Faculty of Social and Information Studies, Gunma University, Maebashi city, 371-8510 Japan. (corresponding author to provide phone: +81-27-220-7440; fax: +81-27-220-7440; e-mail: iwai@si.gunma-u.ac.jp).

H. Masaki is with the Faculty of Social and Information Studies, Gunma University, Maebashi city, 371-8510 Japan. (e-mail: s05601087@ed.si.gunma-u.ac.jp).

Y. Yoshizaki is with the Faculty of Social and Information Studies, Gunma University, Maebashi city, 371-8510 Japan. (e-mail: s05601102@ed.sigunma-u.ac.jp).

designing of this paper. The third section explains the basic design approaches to solve these problems. The fourth section proposes details of supporting methods for critical evaluation and supporting methods for keeping well-ordered discussions. The second and third sections are mainly written by Iwai and the fourth section is mainly written by Masaki and Yoshizaki.

II. Document Evaluation Example and Typical Problems

This section illustrates a specific example of document evaluation and sets out the typical problems that need to be solved in the system designing of this paper.

The example of document evaluation is as follows:

<Example scene>

The main participants are a lecturer and their students in one specific course. Each student independently purchases reference documents for the course, but lacks information concerning which are the most appropriate course materials. The lecturer is obliged to recommend new reference books for the course to the university library, but lacks information which books are most suitable for the students.

Here, the problem of document selection by each student is set as a model of personal decision making, while the problem of document selection by the lecturer is set as a model of group decision making; it is assumed that the main function of the lecturer in this example is to accumulate the preference information of all stakeholders (every student and the lecturer) and to make the final decision of recommendation to the library of the university to which they all belong in order to maximize the potential total utility.

The most common approaches for the students and their lecturer are probably to collect and exchange information in conventional BBS or other open sites managed by large book shops on the Internet, or in a more closed SNS forum where the stakeholders (the students and lecturer) meet.

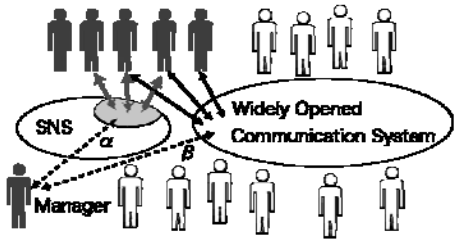

Fig. 1 Collection and Exchange of Information by Using a Computer Network

Fig. 1 is an illustration of such activities of information collection and exchange by the students and their lecturer. Black agents refer to the course stakeholders or students and their lecturer. Here, the lecturer is denoted as the 'Manager.' The other white agents are people who do not belong to the course. The large circle refers to a widely opened communication system like conventional BBS on the Internet or a book shop site that includes review comments by other users. The small circle refers to an SNS where the anonymity of users is limited. The inner gray area of the small circle is a forum where some (but not necessarily all) of the stakeholders gather. For simplicity, arrows to denote participation of outside people are omitted.

First, let us summarize the problem framework from the view point of the lecturer. There are two major alternatives. One is to ask students in the SNS forum (alpha), and the other is to collect information in the open system (beta). In alpha, the information the lecturer obtains may be limited as participants of the forum or students of the lecturer may not input their true opinions that reflect, for example, their critical views concerning the instruction methods of the lecturer (x). (As the appropriateness of a reference book can be connected with the instruction methods of the lecturer, the discussion can extend into an evaluation of the lecturer.)

On the other hand, in beta, the information the lecturer obtains may be again limited, as open systems often contain the noise of useless comments from irrelevant points of view (y). The anonymity of many open systems often leads to context-less, meaningless and disorderly disputes. Emotional and irrational book reviewing comments on many conventional BBS are examples of this problem. A further factor to consider is that even if comments reflect a cool-headed evaluation, as most of the participants in the open systems do not have knowledge of the course, they can not provide a valuable assessment of the appropriateness of a book to the class. For instance, professionals may find no value in some introductory books and leave low evaluation comments in a book shop site while these books may be perfectly adequate for the beginner student. From among so much "noise", it would be a very hard task for a lecturer to find information that is truly valuable to help in choosing their recommendation to the library.

The x mentioned above is the problem of critical evaluation and the y is the problem of orderly maintenance. Notice that there are the same two major alternatives and problems for the students. (For example, the students are also bound by the problem x above as their free information exchange is limited because of the lecturer's participation in the forum.)

The purpose of designing a document evaluation system in this paper is to make a framework for an alternative communication system for the lecturer and students in the above context with which they can efficiently obtain adequate information. The main topic in the designing is how to solve the problems of category x and y.

Now, the illustration of the example of document evaluation is completed.

III. Basic Approaches of System Designing

This section explains our basic design approaches to solve the problems outlined in the previous section.

A. Definition of Participants

First, as a basic framework, the user group of the target system is defined as "all members who belong to the focused organization." As in Fig. 1, the example has two problems: "inclusion of outsiders" (alpha) and "exclusion of some of the members" (beta). Inclusion of outsiders is a problem as it

may result in the inclusion of useless information. Exclusion of some of the members is a problem as it results in the exclusion of needed information. To improve the quality of information exchanged among stakeholders, participation of all members who belong to the organization is the premise of our model.

B. Approach for Critical Evaluation

For supporting critical evaluation, the system utilizes guaranteed anonymity and also employs another mechanism for users to become more involved in the output of critical comments. We implement this mechanism as an application of a collaborative filtering algorithm; a system based on the algorithm is able to naturally collect information concerning the critical evaluation of users while supporting the decision making of each user (See the next section for details).

Here, the definition of participants listed above *A.* is concerned with the method of guaranteeing anonymity. The system must permit every member of an organization to participate, exclude outsiders and guarantee anonymity, which in this case means without conventional individual authentication. This function is implemented by employing a certification framework that requires each participant to use a unique user-name. In this framework, every member of an organization is allocated a unique user-name and password. Each set of user-name and password is given randomly to each member in order to guarantee the anonymity of the user, even to the system manager. The system certifies a user with a user-name as an authorized user by checking the user inputs the correct password for the user-name. In this approach the system makes it possible to guarantee anonymity and certify the user.

C. Approach for Maintaining Orderly Discussion

To ensure orderly discussion, we use an approach employing "the establishment of group rules and the automation of their enforcement." Instead of an autocratic privileged manager with the right to select rules for the system, we propose a framework where through specific procedures participants can select rules together and the enactment of any rule is implemented automatically. "For example, participants can limit the maximum number of comments input to three; in this case, the system prohibits any comment inputting activity after a third comment.

How does this framework maintain an orderly discussion? This concept is based on observation of conflict management in the real world. Conflicts in the real world are normally resolved through legal procedures and rarely lead to physical acts or violence. For example, a man whose family member is harmed by a drunk driver in a car accident would be very upset but rarely try to harm the driver physically. Instead of personally retaliating, he would rely on legal procedures to deal with the problem. As stakeholders are open to using legal procedures to adequately manage conflicts and they understand that such legal procedures "resolve" stressful situations, they can be involved in quickly taking any legal steps instead of engaging in useless accusations.

On the other hand, in many communication systems, only the manager is given the right to delete documents. When a participant starts to engage in anti-social behavior such as insulting other members, and if the manager is not skilled enough, and does not effectively resolve the problem, then the other participants have no other course of action except to retaliate by criticizing the participant in the system. (If the insulted participants could apply rules or introduce a new rule into the system to deal with the problem, they would naturally use the available procedures to redress the situation.) The inability to adequately redress a problem is certainly one of the background reasons for useless slanderous disputes in the conventional BBS system; there is simply no framework for general participants to employ the right to manage the system (Fig.2).

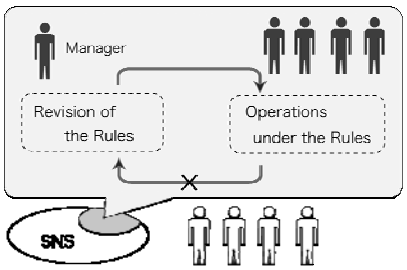

Fig.2 Communication System in Which
Only a Manager Has Control Over the Rules

From this point of view, we employ a simple approach of introducing into our communication system, the real world framework of autonomous group management for maintaining orderly discussion.

Now, the explanation of the three basic design approaches is completed.

IV. Design of Document Evaluation System

This section proposes details of a method to support critical evaluation and a method to support the maintenance of orderly discussion.

A. Design for Supporting Critical Evaluation

For supporting critical evaluation, the system utilizes guaranteed anonymity and also employs a mechanism for users to become more involved in the output of critical comments. This mechanism is implemented as a collaborative filtering system that is linked with A) subsystem designed for inputting negative evaluation information and B) electronic meeting rooms for participants to discuss the suitability of documents to recommend to the library.

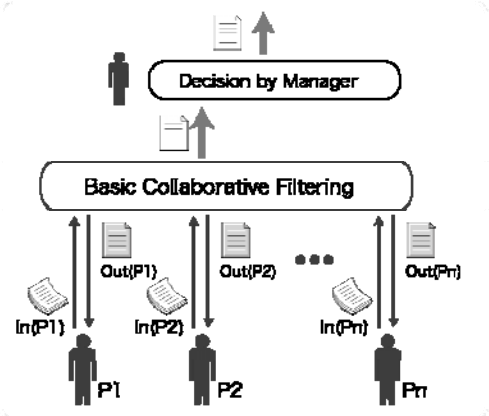

Fig.3 Basic Model of Document Evaluation System

Firstly, the basic system as a skeleton framework without A and B is shown in Fig.3. P1 to Pn denote students and Manager denotes a manager or a lecturer. The students input document evaluation data In(P1) .. In(Pn) and receive recommendation list outputs Out(P1) .. Out(Pn) by the collaborative filtering function. In(P1) is the list of P1's document evaluation data (Each record contains the serial number of the document and a score of P1's evaluation.) Out(P1) is the list of documents which are not evaluated by P1 yet and are expected to be highly evaluated by P1 according to the calculation of the correlation coefficients between P1 and other participants (P2..Pn).

In this system, each student participates basically for their own benefit, which is to create a valuable recommendation list. The manager, however, also receives a list of those documents that are most in demand; this is calculated from the document evaluation data input by the students. The manager can use the information to compile the recommendation to the library.

Here, for document evaluation, the system uses the same binary notation method used in [4], which only reflects whether the user has purchased the document or not. Many collaborative filtering methods require multiple input levels more than 2, as 5 level inputs seen in Grouplens system by Resnick et al.([6]). According to the GroupLens example, based on the input condition, the correlation coefficient of a target user (K) and another specific participant (L) is measured by the following equation.

$$r_{KL} = \frac{Cov(K,L)}{\partial_K \partial_L} = \frac{\sum_i (K_i - \overline{K})(L_i - \overline{L})}{\sqrt{\sum_i (K_i - \overline{K})^2}\sqrt{\sum_i (L_i - \overline{L})^2}}$$

And, using this similarity weight data, prediction on another item (6th item accordingly) for user K is computed as follows.

$$K_{6\,prediction} = \overline{K} + \frac{\sum_{j \in \text{Evaluator}} (J_6 - \overline{J}) \times r_{KJ}}{\sum_j |r_{KJ}|}$$

A many leveled input method, however, has a problem that many participants tend to be reluctant to input their negative evaluation and leave no information of documents that they found unsatisfactory. Besides, to use a many leveled input method, the system requires a relatively high computational specification which is not very common in the office of a general corporation or school that this paper see as the typical organization which would use the system. From these points of view, this paper selected a binary input method that reflects only whether or not the user has made a purchase. Here, the similarity between two sets is simply defined as the proportion of the number of elements shared by the two sets compared to the sum of the elements in both sets.

Now, the approaches for adding A and B to the basic skeleton framework above are as follows.

For adding A, the system employs a function that enables each user to input a list of non-recommended documents that is composed of document data the user considered to be unsuitable for recommendation to others. A non-recommended list is simply called 'Trash box' and is an independent data list that is separate from other document data lists such as In(P1).' (Part A of Fig.4)

The existence of non-recommended list data does not affect the above process of basic collaborative filtering calculations when a participant Px tries to get Out(Px). But, when Out(Px) is output by the system, Px also receives the accumulation of non-recommended data of all users that are 'similar' to Px as information of a negative evaluation. For example, Px may receive information that 30% of other similar participants have already put the top document of Out(Px) into their trash boxes. Px can now use the information to estimate the true value of the document.

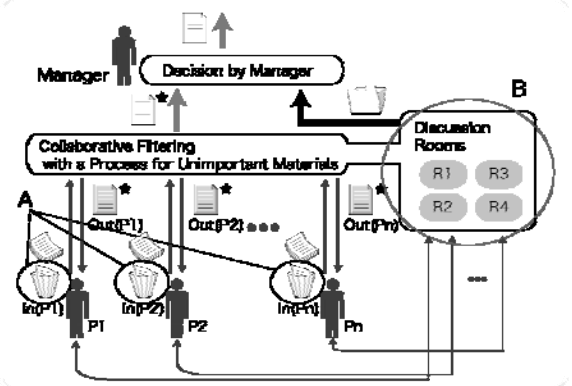

Fig.4 Extension for Supporting Critical Evaluation

Making a non-recommended list is a task completed after finishing normal binary inputting. In the first binary inputting step, the procedure of inputting whether the user has purchased or not is very simple and even a document that has to be given a negative evaluation would be registered with high probability. And for documents registered once, users would have no trouble to give them a negative evaluation. Besides, because of its simple calculation design, this method has little effect on computational speed, while each participant is now able to receive more information (They are able to know the degree of negative evaluation that is shown by the number of other users who felt the document was not suitable for their purpose.)

For adding B, the system employs a function that divides users into groups according to their input values and that motivates users to discuss their evaluation of a specific document. For example, a book that was purchased by most users but 30% of them put in their trash box data list is a difficult example to evaluate. When a manager finds this type of input pattern, they can divide the users who purchased the book into two groups (one for users who just purchased and the other for users who put the document data into their non-recommended list), then invite users to visit the electronic meeting room and encourage them to make comments that expand their view points. (R1..R4 in the Part B of Fig4 are the rooms for this purpose.(Here, the person who leads participants to R1..R4 is the manager. Four is the default number of rooms. However, it can be modified as the number of topics for decision making of document selection for the library increases.).

As the invited users have already expressed their evaluation to the document at this point, it would be easier for them to join the discussion. That is to say, generally, when asked to explain one's attitude to a specific document, many

participants would feel reluctant to join the discussion. Here, we anticipate that it would be easier to join a meeting and express reasons for their previously made document evaluation. If the discussion reveals that for example, the positively evaluated group rate the document as an ideal introductory document and the negatively evaluated group rate it as an inadequate reference document, then this clarification of the apparently contradictory information would be useful in the later decision making of other students and lecturer.

The remaining task is to design a mechanism to maintain arguments that are lead by B as well-ordered discussions among cool-headed participants. This is the next subject of the design.

B. Design for Keeping Ordered Discussion

To support orderly discussion, we employ an approach of "the establishment of group rules and the automation of their enforcement." Instead of endowing a privileged manager with the right of setting rules for the system as in Fig.2, we propose a framework where through specific procedures participants can collaboratively set rules that work as program routines. This revision of rules can be conducted by general participants along with the process of discussion in the meeting rooms R1..R4 (Fig.5).

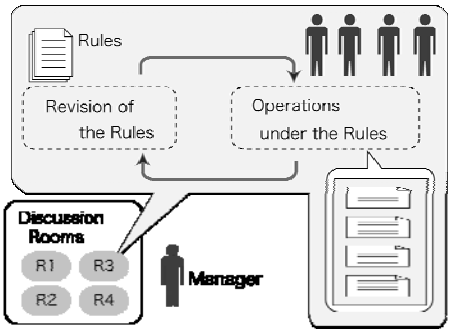

Fig.5 Group Decision of Rules

Fig.6 is an illustration of the procedures for discussion and revision of rules in this framework. Important components are a) a Window for inputting and viewing, b) a Window for making an application to change a rule, c) a List of Rules and d) a Window for voting. Participants can use the Window functions to propose a change to the rules (b) along with using the Window for inputting and viewing (a). The Window for making an application to change a rule (b) is linked to a List of Rules (c) from which a participant can select a target rule and make an application to change it. At this step, a message notifying users of an application to change a rule is shown in the Window for inputting and viewing (a) and each member is allowed to vote in the Window for voting (d). When participants have finished voting, the result is shown in the Window for inputting and viewing (a). If approved, the rule is automatically modified in the List of Rules (c).

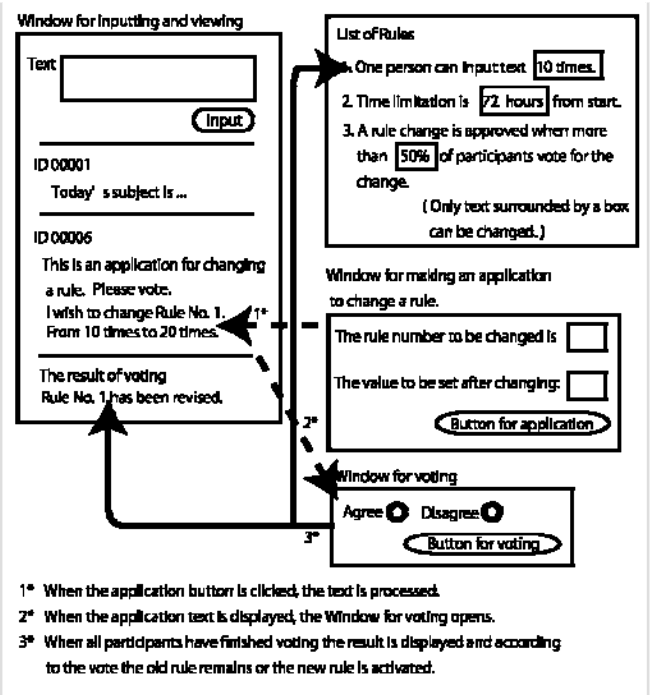

Fig.6 Process of Group Decision of Rules

Typical sample rules: "The maximum number of times for inputting a message is 10," "The time limit for discussion is set at 72 hours from the start time of the discussion," or "Changing a rule needs more than 50% of participants to agree with the new rule." A new rule which proposes the maximum number of times for inputting a message is increased from 10 to 20 is approved automatically when more than 50% of the participants agree with this proposal. Notice that the critical value of more than 50% can itself be a subject of revision. (For example, a proposal to change more than 50% to more than 60% is approved by more than 50% agreement. If a proposal to return the rate to more than 50% is made later, then it needs more than 60% agreement for approval.)

This design makes use of conflict management models found in the real world where procedures to deal with problems are open to participants. That is to say, when the behavior of a participant falls below certain standards, e.g. insulting other members, the other members can apply sanctions by using already established rules or propose new rules for the system to deal with the problem. This framework is expected to help participants avoid exchanges of useless insults.

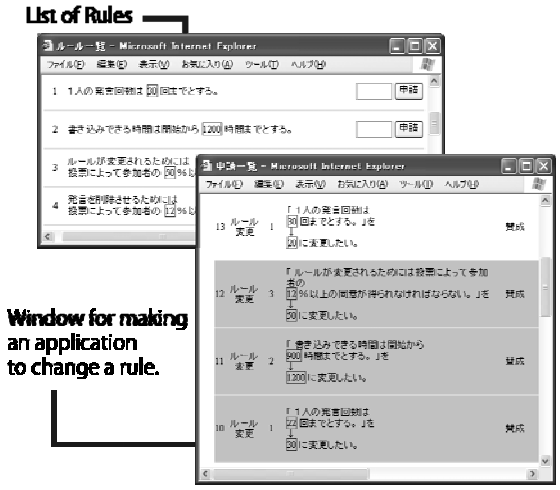

Fig.7 A Prototype System (Japanese Language)

We have already implemented a prototype of the system (Fig.7) and performed experiments; although these experiments are simple, we learned that even by the simple introduction of an input limitation rule, user behavior can be more serious, and that a user can accept the rule partly because of the autonomous feature of rule setting. Verification of the effectiveness of the function needs more detailed design and practical experiments. The design above is based on a basic mechanism for maintaining orderly human interactions in the real world, and it is reasonable to expect that it would have some level of effectiveness.

Now, this section has proposed details of methods to support critical evaluation and orderly discussion.

V. Concluding Remarks

This paper presented a new design of a system that can support users to output critical evaluation comments and can also maintain autonomous order at the same time; the system would be useful in such fields as book-reviewing where communication of critical evaluation is essential. We demonstrated the design as a framework of a document evaluation system for personal and group decision making for document selection.

Our design for supporting critical evaluation includes guaranteed anonymity and utilization of a collaborative filtering algorithm. The most original point in our design for supporting critical evaluation is the framework which naturally links individual activity to produce a recommendation list for an individual with the greater activity of group decision making for recommending a document to a library.

Using archived data on the Internet, we have implemented a prototype system which includes data for 5,089 volunteer users and data for 335,132 documents that belong to the users (the data contributes to increase the reliability of the calculation of a recommendation list). For inputting non-recommended lists, we employed two types of implementation. In the negative model, a user can label “not-recommendable” to each item when inputting document data and in the reverse model, a user can label “favorite.” In the reverse model, any document that is not labeled is regarded negatively. One of the concerns in evaluation of the system would be that introducing a non-recommendation list may make the input interface more complex and decrease the number of document data inputs. However, in our simple experiment that involved 167 students, in terms of frequency of data input we found that both of the models were preferred rather more by users than the original model without label input interface. In terms of collecting negative evaluation information, the reverse model is more effective than the negative one. The detail analysis is to be focused on later.

Our design for maintaining orderly discussion is very original and there are few similar projects. To check its effectiveness level, a prototype system with a more detailed design needs to be developed and experiments undertaken. These are the tasks for the next studies.

Recent study related to the framework of maintaining well-ordered discussion in this study includes [1], [2] and [3] that visualize discussion structure for the purpose of supporting participants or facilitators. It is characteristic of this study not to focus on the discussion structure but to realize self-management by the group for maintaining and changing the rules to promote orderly discussion.

In terms of self-management by the group, this study is also related to [5] that proposed a framework in which a possibly illegal document in a communication system is concealed under the agreement of participants. The framework is similar to our approach, especially in the point that the agreement of participants is arrived at through a voting procedure. Our framework of self-management by the group, however, deals with the discussion rules instead of the documents themselves. From this point of view, we expect the framework of this study has more potential in actual application.

References

[1] M. Horita, T. Enoto, N. Iwahashi, “A Pluralist Approach to Visualisation of Policy Discourse: Argumentation Support as Socio-Technical Systems”, *SOCIOTECHNICA*, Vol.1, 2003, pp.67-76.

[2] K. Ito, R. Samejima, Koji Matsui, Hidekazu Yoshikawa, “A Trial Study for School Education by Using Computerized Group Discussion Support System: Proposal for Utilization of ‘ERDE’ in High School and Graduate School”, *Journal of the Japan Association for Social Informatics*, Vol.16, No.2, 2004, pp.5-19.

[3] Y. Makino, N. Pan, “Development of Web Interface for Structuring Argumentation: Facilitationg Argumentation with Message Construction Cross” *JSET08-1(Japan Society for Educational Technology)*, 2008, pp171-174.

[4] T. Masui, Moderate Group Communication on Shared Bookshelves, *Proceedings of Interaction 2005* (Infromation Processing Society of Japan), 2005, pp135-142.

[5] K. Sado, A. Iwai, “Sankasya no Goui ni motozuku Bunsho Inpei Kikou no Jissoujou no Ronten”, *Research on the Japan Model Construction of Electronic Democrary* (Report of the Grant-in-Aid for Scientific Research(C) No.15500154 edited by Yoshinori Tomiyama), 2006, pp53-63.

[6] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm and J. Riedl, “GroupLens: An Open Architecture for Collaborative Filtering of Netnews”, *Proceedings of the 1994 Computer Supported Collaborative Work Conference*, 1994, pp.175-186.