# Speaker Choice Method based on Multi-armed Bandit Algorithm for Online Discussions

Mio Kurii
Tokyo University of
Agriculture and Technology
Tokyo, Japan
Email: kurii@katfuji.lab.tuat.ac.jp

Katsuhide Fujita
Tokyo University of
Agriculture and Technology
Tokyo, Japan
Email: fujita@katfuji.lab.tuat.ac.jp

***Abstract*—This paper proposes an application of the multi-armed bandit algorithm to online discussions. We assume a speaker choice in discussion by a facilitator as a multi-armed bandit problem : Each participant is considered as an arm of a slot machine, and a facilitator as a player. The facilitator's behavior when they select one participant can be considered to be equivalent to the behavior of a player who selects one slot machine and plays it in the multi-armed bandit problem. As a reward of slot machines, we define a "discussion score" to evaluate each post. In addition, in order to consider conflict between participants in a discussion, our method classifies the participants into groups and determines the next speaker based on clustering results. We demonstrate that our method can select participants who posted good ideas and opinions and promote participants to engage other participants by using questionnaires.**
**keywords : Multi-Armed Bandit Problem, Decision Support System, Automated Facilitator**

## I. Introduction

Online discussions are becoming increasingly prevalent owing to the popularity of smartphones and social networking services. Online discussions have many advantages over other forms of communication. For example, it is possible to conduct meetings regardless of the users' locations and to record the history of conversations. Thus, it is possible that the use of online discussions is going to increase on a large scale in the coming years. However, this form of remote meeting method also has some disadvantages. Users have difficulties reading and conveying indirect communications, including gestures, nervous habits, room tension, and eye contact. Therefore, problems such as conversation delays and miscommunications can easily occur. For important discussions, people usually prefer face-to-face meetings over online discussions because the problems mentioned above may lead to misunderstandings and negative side effects. Therefore, the development of online discussion support systems could lead to finding solutions that can solve the problems.

Usually, the functions of discussion facilitators are as follows:

- providing topics and questions
- making an atmosphere where everyone can talk openly
- preventing a discussion from going on a tangent from the main subject
- summarizing the manner in which the discussion was done and giving participants feedback
- determining the next speaker

This paper focuses on the action in which subsequent speakers can be determined using a facilitator. The participant determination system is an important for deciding who the facilitator will select as the next speaker.

This paper proposes a method to determine subsequent speakers in an online discussion using a multi-armed bandit algorithm. This is one of the important tasks assigned to the facilitators. Bandit algorithms can be applied to speaker determination systems by considering each participant as an arm of a slot machine and a facilitator as a player. We define a discussion score to evaluate each post; this is considered to be equivalent to a reward based on the slot machine metaphor. The discussion score of each post is defined based on the following three assessment measures: (1) Whether the post helps to settle a discussion or not (2) How interested the other participants are in the post (3) The intention of the post. To analyze the conflict among participants, our method classifies the participants into groups and determines the next speaker based on clustering results. We demonstrate that our method can select participants that post good ideas and opinions and promote the posts of participants to encourage participation using questionnaires.

The remainder of this paper is organized as follows. First, we discuss related research and explain Multi-armed Bandit Problem and Algorithms. Next, we propose the subsequent speaker determination method based on the multi-armed bandit algorithm. Then, we demonstrate our experimental results. Finally, we present our conclusions.

## II. Related Work

### A. Automated Facilitators in Online Discussions

Although there is no existing research that focuses on "choosing the next speaker" (based on this literature review), the importance of facilitators in online discussions is becoming widely known. Yuuki et al. [1] proposed a facilitator agent for online meetings. The agent has the following two main functions: to generate diverging discussions (to gather many ideas/opinions) and to generate convergent discussions (to settle discussions). The method comprises the following three steps. (1) Analysis of the participants statements (2) Determination of the approach to intervene in the discussion (3) Speech generation. The way it intervene in the discussion

contains “Determining the next speaker.” When a participant has not posted anything for a certain time, the agent asked his/her opinion.
Ito et al. [2] proposed an online discussion platform with a facilitator support system called “COLLAGREE.” One example of how this system functions is that it judges whether each post is positive or negative using an indicating function that shows significant keywords. It provides the templates for efficient facilitation to help human facilitators to function properly. These templates are generated based on the human facilitators statements.

### B. Application of Multi-armed bandit problem

Multi-armed bandit algorithm is applied to various problems such as game tree searches[4], clinical trials[5], and interface design optimizations[6].
Web advertising system is one of the famous examples of an application of the Multi-armed bandit approach. A study of Li et al. is one similar example. They apply bandit algorithms to personalized web-based services such as news article recommendation[3].

In addition, this paper is an extension of our paper[7] by adding explanation of our method and more details of experimental results and discussions.

## III. Multi-armed Bandit Problem and Algorithms

This section shows an abstract of the problem and its algorithm based on Kuleshov and Precup’s study[8].

### A. Multi-armed bandit problem

The multi-armed bandit problem solved one of the optimization problems. There are $K$ slots, and a player chooses one of them and plays. The objective is to maximize the sum of rewards of $T$ times plays. The result from each slot action is either a hit or a miss($0$ $or$ $1$). The expectation for each slot action is different. The player cannot know its true value. Under the above conditions, the player has to select the machine to play on, the number of times to play, and the order in which the activities are performed. Assume that there are five slots ($N = 5$) and a player will play the slots 100 times ($T = 100$). The player wants to play 100 times with the slot with the greatest payout, but the player has no information about the slots and cannot select the machine that has the best payout. Then, the player decides to play each slot $n$ times and subsequently play with the arm that has the highest average reward for the rest of the $(T - nK)$times.If the value of $n$ is set to significantly a small value, it is difficult to determine the best slot because the rest of $nK$ times may not be sufficiently large to determine it. However, if the value of $n$ is too large, the player can barely make any profit because $(T - nK)$ times may be too small. A player has to perform both exploration and exploitation, which are in a trade-off relationship, at the same time. The general descriptive family of multi-armed bandit problems is called bandit problems. The bandit algorithms are not only used for slot machine problems but also are used for many types of optimization problems that need to analyze trade-offs between exploration and exploitation. Examples are shown in section II.

### B. Upper Confidence Bound Policy

The Upper Confidence Bound (UCB) policy is one of the well-known multi-armed bandit policies. We adopted this policy as the proposed method in this research. In this policy, a player tries each slot once equally at first. Next, the algorithm calculates the UCB score of each slot and selects the slot that has the largest value each time. The UCB score is defined below:

$$\bar{\mu_i}(t) = \hat{\mu_i}(t) + \sqrt{\frac{\log t}{2N_i(t)}} \tag{1}$$

This is the sum of the average of the rewards($\hat{\mu_i}(t)$) and the correction term($\sqrt{\frac{\log t}{2N_i(t)}}$). $N_i(t)$ means the total number of times that arm $i$ has been selected at the time $t$. The smaller the value of $N_i$, the greater the correction term; this means that the slots that have a small number of samples are easily selected, even in a case in which the average of the samples is small.

## IV. Subsequent Speaker Determination Method based on the Multi-armed Bandit Algorithm

### A. Applying the Bandit Algorithm to Speaker Determination

In this paper, we assume that speaker determination can be represented as a bandit problem. The basis of this assumption is shown below. To make a discussion active and productive, the most assertive people need to have the ability to discuss and speak openly so that the others are stimulated. However, the assertiveness of the participants in discussions depends on factors such as the main theme of discussion, the type of discussion, and the relationship between each participant. It is not easy to estimate the assertiveness of the participant. It gradually becomes clear during discussions. These conditions are similar to those in the multi-armed bandit problem. Any policies using the multi-armed bandit problem should have a strategy based on the information gained from past trials. From the above basis, we apply a bandit algorithm to speaker determination. We apply the UCB policy to speaker determination to enable a facilitator to select a suitable speaker based on the posts of the past.
The details of the action in which the bandit algorithm is applied to speaker determination are as follows. The facilitator’s behavior when they select a participant and urge the participant to speak can be considered to be equivalent to the behavior of a player who selects one slot and plays it in the multi-armed bandit problem. Accordingly, the participants of a discussion can be regarded as slots. The reward of a slot action is either a hit or a miss (0 or 1). In speaker determination, we defined a term called the discussion score that reflects the influence of the post posted by a selected participant. We consider this score to be equivalent to a reward in the multi-armed bandit problem. Figure 1 shows the overview of it.

### B. Discussion Score

A discussion score reflects the influence of a post posted by a selected participant. A discussion scores $f_p$ of post $p$ is defined based on three points as follows: The first point is **whether the post helps to solve the issues in a discussion or not ($f_1$).** In online discussions, it tends to be particularly

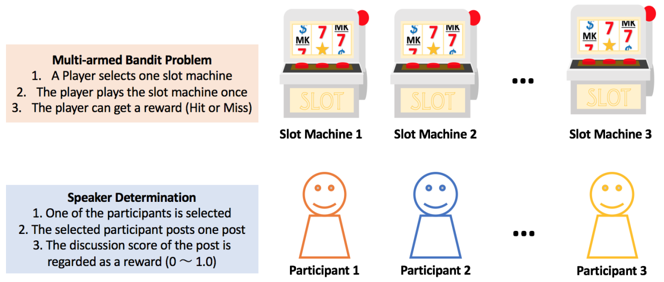


Fig. 1. Overview of Applying the Bandit Algorithm to Speaker Determination

TABLE I. DEFINITION OF CATEGORY TAGS

| category tags | definition |
|---|---|
| #proposal | proposes one's new opinions, ideas, or thought |
| #explanation | presents information or details of something |
| #supporting | includes an opinion which supports another post |
| #confronting | includes an opinion which refutes another post |
| #question | asks a question about another post or to the whole discussion |
| #answer | answers to a question |
| #etc | anything that does not fit any of the above |

difficult to put each opinion together and make an agreement compared to face-to-face discussions. We assume that the discussion system has "Agree" and "Disagree" buttons so that each participant can evaluate each post intuitively an independently. Although general social medias tend to have only functions of favorable intention (ex. "Like" functions on Instagram), out discussion system also has "Disagree" function. It enables people to view the ratio between "Agrees" and "Disagrees." and make it smooth to reach a conclusion. Posts that gain much support from participants can be considered to help in solving the issues in a discussion. $f_1$ of post $p$ based on this concept is defined as follows:

$$f_1(p) = \frac{N_{agree}(p) - N_{disagree}(p)}{K} \tag{2}$$

$N_{agree}(p)$ means the number of "Agree" that post $p$ gets, so as $N_{disagree}(p)$. $K$ is the number of participants.

The second point is **how interested are the other participants in the post ($f_2$).** A post that attracts interest from participants can make a discussion active and influence it positively. It can be measured by the number of replies that other participants sent to the post. Therefore, $f_2$ of post $p$ based on this concept is defined as follows:

$$f_2(p) = \frac{N_{reply}}{K} \tag{3}$$

$N_{reply}(p)$ means the number of replies to the post $p$.

The third point is **the intention of the post ($f_3$).** The type of post is a very important factor in post categorization. Some online discussion systems have a category tagging function to reflect the statement intention. In this study, seven types of categories are defined in a table I, based on Kotani et al[9].

Posts categorized as "#proposals" are considered to have a good influence on a discussion; these could include new propositions about a main theme of a discussion and new ideas on how to progress with the discussion. Additionally, posts

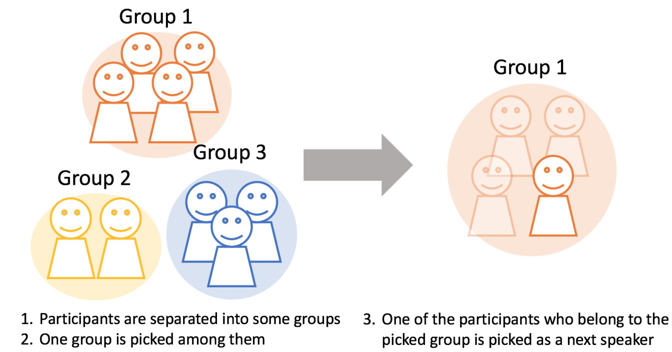


Fig. 2. Overview of Double Bandit

categorized as "#explanations" are considered to increase the productivity of a discussion. For example, objective facts, evidence in support of one's claims, and so on. Since these two above are considered to be important category tags, we reflect it in the discussion score. We define $f_3(p) = 1$ if the category tag of post $p$ is a"#proposal" or an"#explanation", otherwise $f_3(p) = 0$.

The above three scores are weighted based on each of their importance. $f_1$ can be considered to have the highest importance because it directly reflects a participant's preference. $f_2$ has the highest importance because it depends on how active the discussion is. $f_3$ is the lowest.

Using this rank, they are weighed as follows:

$$f_1(p) : f_2(p) : f_3(p) = 3 : 2 : 1 \tag{4}$$

Finally the discussion score $f(p)$ of post $p$ is defined as follows.

$$f(p) = \frac{3}{6}f_1(p) + \frac{2}{6}f_2(p) + \frac{1}{6}f_3(p) \tag{5}$$

### C. Clustering using Bipartite Graph

Omoto[10] revealed the importance of relationships between participants, particularly conflict/cooperative relationships to make an agreement. Facilitators need to consider it too. Specifically, it is desirable that the participants are separated into groups based on their opinions or preferences, and a facilitator should be careful not to show bias in the number of total statements permitted for each group. When the bandit algorithm is applied to speaker determination, the relationships between participants are not considered. To solve this problem, we propose the use of a "two-step bandit application." A "two-step" approach means that the participants are divided into groups, and they select one group first and then select one participant from the group. Here, we apply the bandit algorithm (UCB policy) for both group determination and speaker determination. Figure 2 shows the concept of the "double bandit method." To apply the UCB policy to group determination, we consider each group to be a slot. To calculate the UCB score of group $G$, the following two values are necessary.

1) The average discussion scores from all the posts by the participants belonging to group $G$

2) The total number of posts from participants belonging to group $G$

To divide the participants into appropriate groups, participants with similar opinions should be put in the same group, while participants with contrasting opinions should be separated from each other. We applied the clustering method using the bipartite graph that was used by Nakahara et al.[11] for tweet clustering. Then, the participants are clustered based on the "Agree" and "Disagree" overlap degree of each post between participants. This paper assumes the union of posts in which user $U_i$ puts "Agree" is $P_a(i)$ and user $U_i$ puts "Disagree" is $P_d(i)$, The overlap degree of the agreement between user$A$ and user$B$ is calculated by the ratio of $P_a(A) \cap P_a(B)$ to $P_a(A) \cup P_a(B)$ (Jaccard similarity coefficient); this gives the overlap degree of disagreement. Participants with a large overlap degree value are considered to have similar opinions or preferences. This paper defines a similarity value $O_{ij}$ between user$U_i$ and user$U_j$ as follows.

$$O_{ij} = \frac{|P_a(A) \cap P_a(B)| + |P_d(A) \cap P_d(B)|}{|P_a(A) \cup P_a(B)| + |P_d(A) \cup P_d(B)|} \tag{6}$$

The next step is the construction of a network graph. A network graph is constructed based on the similarity values between participants. Each node represents one participant. If the similarity of two participants is larger than $th$, their nodes are linked. After this procedure, $best\_partition$ is applied to the network graph. $best\_partition$ is a clustering function included in $community$(a library of Python). We set the value of $th$ to 0.45 based on a preliminary simulation using existing discussion data.

## V. Experiments

The purpose of the experiment was to evaluate the validity and the usefulness of our method by applying it to real online discussions.

### A. Experimental Settings

We conducted three discussions with 21 subjects and three groups with seven people in each group. All discussions were held on an online discussion system we created.
Each discussion lasts 75 mins, which are separated into the former (25mins) and the latter (50mins) parts. The former part is an open discussion. The latter part takes place under the following rule: Each of the participants picked by a facilitator can post once (1 post only)." The facilitator conducts the experiment, and determines the next speaker among the participants according to the speaker determination system. After the selected participant posts, or two minutes pass after the participant was selected, then the facilitator selects the next speaker. This process is repeated until the end of a discussion. We do not tell the participants whether the facilitator selected the next speaker each time based on our proposed method or using a baseline. Before the experiments, the facilitator informs the participants about the goals of the discussion: "presenting as many ideas as possible" and "reaching a conclusion that every participant can agree upon."
We use two speaker selection determination methods; one of them determined the next speaker based on the proposed method, and the other determined the next speaker randomly.

TABLE II. The list of Discussion's Conditions

| | Group 1 | Group 2 | Group 3 |
|---|---|---|---|
| Discussion 1 | Topic 1<br>Baseline Method | Topic 1<br>Proposed Method | Topic 2<br>Baseline Method |
| Discussion 2 | Topic 2<br>Proposed Method | Topic 2<br>Baseline Method | Topic 1<br>Proposed Method |

Two different speaker determination approaches were used in separate experiments to compare them. Table II shows the topic and method was used in each discussion. We added two rules so that the discussion appeared to be natural.

- Both the proposed method and the baseline did not select the same participant two times in a row.
- In the UCB policy, the slot with the highest UCB score should be selected. In this experiment, we turned it into a stochastic selection system. The next speaker was selected randomly based on the expectation of each participant, which was in direct proportion to their UCB score. For example, if the ratio of the UCB score between the participants $u_1$ and $u_2$ is $5:1$, the probability that the system picks $u_1$ is five times as high as that of $u_2$.

Questionnaires about the experiment were conducted to evaluate our proposed method.

1) **There were many good ideas and opinions in the latter part of the discussion**
   a) Strongly disagree
   b) Disagree
   c) Neither agree nor disagree
   d) Agree
   e) Strongly agree
2) **There were only limited opinions and ideas in the latter part of the discussion**
   a) Strongly disagree
   b) disagree
   c) Neither agree nor disagree
   d) Agree
   e) Strongly agree
3) **The other participants' ideas and opinions contributed to your final opinion on the topic during the discussion**
   a) Strongly disagree
   b) disagree
   c) Neither agree nor disagree
   d) Agree
   e) Strongly agree
4) **Were the speakers appropriately determined by the facilitator?**
   a) Strongly inappropriate and unnatural
   b) Inappropriate and unnatural
   c) felt nothing, unconscious
   d) Appropriate and natural
   e) Strongly appropriate and natural
5) **Evaluate the other participants' discussion abilities**
   a) Very poor
   b) Poor

TABLE III. TOTAL NUMBER OF POSTS IN EACH DISCUSSION

| | Group 1 | Group 2 | Group 3 |
|---|---|---|---|
| Discussion 1 | 70 (42:28) | 77 (44:33) | 119 (85:34) |
| Discussion 2 | 63 (36:27) | 76 (47:29) | 127 (97:30) |

TABLE IV. TOTAL NUMBER OF POSTS OF EACH CATEGORY IN EACH DISCUSSION

| | #prop. | #exp. | #supp. | #conf. | #ques. | #ans. | #etc |
|---|---|---|---|---|---|---|---|
| Group1 Discussion1 | 30 | 3 | 16 | 5 | 3 | 3 | 10 |
| Group1 Discussion2 | 18 | 2 | 13 | 6 | 6 | 6 | 12 |
| Group2 Discussion1 | 34 | 4 | 16 | 5 | 5 | 4 | 9 |
| Group2 Discussion2 | 39 | 6 | 21 | 1 | 0 | 2 | 7 |
| Group3 Discussion1 | 41 | 3 | 24 | 9 | 3 | 2 | 37 |
| Group3 Discussion2 | 51 | 3 | 14 | 11 | 4 | 3 | 41 |

c) Neither high nor poor
d) High
e) Very high

### B. Results

Table III shows the totals from each discussion (The former and the latter). In all of them, the former parts had larger numbers of posts than the latter parts. Table IV shows the total number of posts of each category. #propsal and #supporting are more frequently used compared to the other category tags. Table V shows the total number of Agrees and Disagrees. Apparently the participants tend to use Agree function more than Disagree function.

1) **There were many good ideas and opinions in the latter part of the discussion**
Table VI shows the results of 21 subjects. The percentage of those from our proposed method who answered "Agree" (includes (d) and (e)) is 72%, even though the baseline is 62%. Each group's result also shows consistent results (the proposed method's percentage was larger than that of the baseline method). Thus, the proposed method can successfully select participants who posted good ideas and opinions. The UCB policy tends to select a slot which has not been picked for a sufficient number of times before. Participants who have not posted as much in a former part tends to be picked more than those who had posted a lot. In the latter part, those participants often posted interesting ideas and opinions from a new perspective.

TABLE V. TOTAL NUMBER OF AGREES AND DISAGREES OF EACH DISCUSSION

| | Number of Agree | Number of Disagree |
|---|---|---|
| Group1 Discussion1 | 142 | 6 |
| Group1 Discussion2 | 125 | 19 |
| Group2 Discussion1 | 212 | 42 |
| Group2 Discussion2 | 229 | 27 |
| Group3 Discussion1 | 167 | 28 |
| Group3 Discussion2 | 227 | 36 |

TABLE VI. THERE WERE MANY GOOD IDEAS AND OPINIONS IN THE LATTER PART OF THE DISCUSSION

| | (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|---|
| Baseline Method | 0 (0.0%) | 1 (4.8%) | 1 (4.8%) | 4 (19.0%) | 1 (4.8%) |
| Proposed Method | 0 (0.0%) | 2 (9.5%) | 2 (9.5%) | 2 (9.5%) | 1 (4.8%) |

TABLE VII. THERE WERE ONLY LIMITED OPINIONS AND IDEAS IN THE LATTER PART OF THE DISCUSSION

| | (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|---|
| Baseline Method | 0 (0.0%) | 2 (9.5%) | 4 (19.0%) | 13 (61.9%) | 2 (9.5%) |
| Proposed Method | 1 (4.8%) | 6 (28.6%) | 1 (4.8%) | 11 (52.4%) | 2 (9.5%) |

2) **There were only limited opinions and ideas in the latter part of the discussion**
This question is asked to confirm the effect of the clustering and the double bandit method. Table VII shows the results of the question. The percentage of who answered "Disagree" (including (a) and (b)) in the proposed method was 10% larger than that of the baseline method (34%). The results of each group were also the same (the percentage of the proposed method was larger than that of the baseline). Therefore, the proposed method prevented the discussions from having only similar ideas and opinions. In addition, the proposed method promoted diverse ideas and opinions.

3) **The other participants' opinions contributed to your final opinion on the topic**
This question was asked to confirm whether the proposed method can contribute to the productivity of the discussions. Table VIII shows the results. Exchanging ideas and opinions actively between participants is one of the most important purpose of discussions. This question can be considered to be an efficient approach to measure it. Using the proposed method, the percentage of those who answered "Agree" (includes (d) and (e)) increased from 67% to 86%, compared to the baseline method. Owing to the above results, the proposed method inspired the participants to engage each other. These results are considered to be strongly related to the results of the questions above.

4) **Were the speakers determined by the facilitator appropriately?**
Table IX shows the results. By using the proposed method, the percentage of respondents who answered "Inappropriate" (includes (a) and (b)) increased from

TABLE VIII. THE OTHER PARTICIPANTS' IDEAS AND OPINIONS CONTRIBUTED TO YOUR FINAL OPINION ON THE TOPIC THROUGH THE DISCUSSION

| | (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|---|
| Baseline Method | 1 (4.8%) | 2 (9.5%) | 4 (19.0%) | 6 (28.6%) | 8 (38.1%) |
| Proposed Method | 0 (0.0%) | 3 (14.3%) | 0 (0.0%) | 9 (42.9%) | 9 (42.9%) |

TABLE IX. WAS THE SPEAKER DETERMINATION BY THE FACILITATOR APPROPRIATE?

| | (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|---|
| Baseline Method | 0 (0.0%) | 2 (9.5%) | 14 (66.7%) | 4 (19.0%) | 1 (4.8%) |
| Proposed Method | 0 (0.0%) | 7 (33.3%) | 9 (42.9%) | 5 (23.8%) | 0 (0.0%) |

9% to 33%, compared to the baseline method. Even though the proposed method had a good effect, the number of people who felt that the method was inappropriate or unnatural was larger for the proposed method than that for the baseline. One of the reasons for this result was that few participants answered that the number of speakers selected seemed to be either too many and frequent, or too few in the proposed method. Since the baseline method selected the next speaker randomly, the total number of selected participants could have been evenly distributed. However, the proposed method could be considered to be somewhat "picky" and the total number of selected participants could have been uneven. Since it might effect participants' feelings of satisfaction, we need to find a proper solution for it.

5) **Evaluate the other participants' discussion abilities**
The options for this question comprised five rankings. We defined "Very poor" as 0 and "Very high" as 5 and quantified each participant's discussion ability by adding the total amount of answers. To confirm the validity, we calculated (1)The correlation coefficient $\rho_1$ between "the discussion abilities" and "the average of the discussion score" of each subject, and (2)The correlation coefficient$\rho_2$ between the discussion abilities and "the total of the discussion score" of each subject. These values became $\rho_1 = 0.037$ and $\rho_2 = 0.349$. Although both of them were positive values, the correlations were not very strong. Thus, improvement of the definition of the discussion scores is a possible future research option.

We can summarize the experimental results as follows:

- The proposed method selected suitable participants who posted good ideas and opinions.
- The proposed method prevented discussions from having only similar ideas and opinions.
- The proposed method promoted participants so that they could stimulate each other.
- The infrequent number of times each participant was selected could make them feel unfair.

## VI. CONCLUSION

This paper proposed an application of the multi-armed bandit algorithm to online discussions. We conducted real discussion experiments to confirm the validity and the usefulness of the method. Based on the experimental results, we confirmed that our proposed method can positively influence discussions; for example, it increased good ideas and opinions and prevented the discussions from having only similar ideas and opinions. One of the possible future research options is to decide the most effective timing method when allowing someone to speak. The system should be able to identify the situations and does conduct speaker determination automatically. Our task is to find a way to implement a discussion state recognition model that recognizes particular states which need a facilitator's intervention.
Another possible future research is to apply the Multi-armed Bandit Algorithm to another human communication's problems.
We obtained interesting dialogues of online discussions from the experiments. Those data can be utilized to analyze participants' behaviors in a discussion and help our future research.

## ACKNOWLEDGMENT

This work was supported by JST CREST Grant Number JPMJCR15E1, Japan.

## REFERENCES

[1] D. Yuuki, S. Matsuyoshi, and A. Utsumi, "A facilitator agent for opinion diversity and discussion convergence in an online text chat meeting," *2017 Information Processing Society of Japan*, vol. 2017-NL-231, no. 17, pp. 1–10, 2017.

[2] T. Ito, S. Fukamachi, K. Tanaka, T. Ito, and E. Hideshima, "Veri cation and evaluation of the supporting system of consensus building focus on facilitator verification and evaluation of the supporting system of consensus building focus on facilitator —a case study on product development of office furniture," *BULLETIN OF JSSD*, vol. 62, no. 4, pp. 67–76, 2015.

[3] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proceedings of the 19th International Conference on World Wide Web*, 2010, pp. 661–670.

[4] S. Ontanon, "Bandit algorithms in game tree search: Application to computer renju," *Journal of Artificial Intelligence Research 58*, pp. 665–702, 2017.

[5] S. S. Villara, J. Bowden, and J. Wason, "Multi-armed bandit models for the optimal design of clinical trials: Benefits and challenges," *Statistical Science, 2015, Vol. 30, No. 2*, p. 199215, 2015.

[6] J. D. Lomas, J. Forlizzi, N. Poonwala, N. Patel, S. Shodhan, K. Patel, K. Koedinger, and E. Brunskill, "Interface design optimization as a multi-armed bandit problem," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 4142–4153.

[7] K. Mio and K. Fujita, "A subsequent speaker selection method for online discussions based on the multi-armed bandit algorithm," *Pacific Rim International Conference on Artificial Intelligence 2018*, pp. 404–411, 2018.

[8] V. Kuleshov and D. Precup, "Algorithms for the multi-armed bandit problem," *Journal of Machine Learning Research*, vol. 1, pp. 1–48, 2000.

[9] T. Kotani, K. Seki, T. Matsui, and T. Okamoto, "Development of discussion supporting system based on the"value of favorable words' influence"," *The Japanese Society for Artificial Intelligence*, vol. 19, no. 2, pp. 95–104, 2004.

[10] Y. Ohmoto, Y. Toda, K. Ueda, and T. Nishida, "Analyses of the facilitating behavior by using participant ' s agreement and nonverbal behavior," *2011 Information Processing Society of Japan*, pp. 3659–3670, 2011.

[11] T. Nakahara, A. Ouchi, T. Uno, and Y. Hamuro, "Extracted opinions from twitter using bipartite graph polishing," *The 30th Annual Conference of the Japanese Society for Artificial Intelligence*, pp. 1–4, 2016.