Translation

INFORMATION AND TECHNOLOGY IN EDUCATION AND LEARNING, Vol. 3, No. 1 2023
https://doi.org/10.12937/itel.3.1.Trans.p003

A Recommendation-type Dialogue System Responding to Potential Requests in Consideration of Personal Attributes

Yosuke Seki* **and Yoshihiko Ueno***

(The original Japanese article was published in Japan Journal of Educational Technology, Vol. 45, No. 1, 2021, pp. 103–112.)

Abstract A dialogue system that responds mechanically to given queries can be used for public relations for prospective university students. However, as a user-specific problem, collecting information from first-time entrances and passive examinations is difficult, and a conventional dialogue system may not collect sufficient information. Therefore, we developed a recommendation-type dialogue system that supports information gathering by evoking potential user requests. Based on the preference analysis results using previous and real-time query histories and recommended information from universities, information in the form of queries is recommended to users when considering their personal attributes. We evaluated the presence or absence of a recommendation function and implemented an introduction evaluation targeting public relations for those who wish to enroll in a university. The average number of queries increases by evoking potential requests, and the information gathering of the user is supported.

Keywords dialogue system, information recommendation, collaborative filtering, public relation, entrance examination in universities

1. Introduction

As the population of 18-year-olds declines, and we enter an era of universal university admissions, universities must secure better applicants. In addition, owing to a reform of the high school-to-university connection, a multifaceted and comprehensive evaluation of the three elements of academic capabilities is required, and various reforms are being promoted at national universities. Against this backdrop, emphasis has been placed on public relations to disseminate information on university characteristics, new selection methods, job performance, and living environments.

Public relations of a university can be broadly divided into the distribution and publication of public relations media (university guides, websites, etc.) and face-to-face communication through counseling, information, and other types of sessions (Fig. 1). However, the former may be overlooked owing to scattered information, whereas the latter is constrained by the number of participants due to geographical and scheduling issues. Therefore, we must establish new means of public relations to solve these problems, allowing the required information to reach those who wish to pursue higher education.

Recently, dialogue systems that respond mechanically to given queries have attracted attention for automating inquiry responses and reducing human resources. For instance, governments and companies have actively introduced such technologies, including garbage sorting and tourist guide services. This technology has reached a practical stage; therefore, dialogue systems can also be used for public relations for those who wish to pursue higher education. In particular, because we can obtain specific information that the user is interested in, the dialogue system is considered effective in resolving explicit user requests.

However, most of those who want to pursue higher education are high school students. Not knowing what questions to ask, taking entrance exams for the first

Figure 1. Major public relations at universities

* Research Center for Higher Education, Tokushima University, Japan

time, the use of passive examinations and complex entrance examination structures, and the diversification of universities have made it difficult for such students to gather information. Therefore, although there is reference information in the dialogue system, we may fail to solve the user's latent request, such as not asking questions, because the content is confusing. For example, it is possible that important information, such as characteristic educational content and admission policies (AP), will not be collected by students when they decide where to enroll or take entrance examinations. Owing to the insufficient information collection, there are concerns about mismatches after admission and unfairness in examinations.

Therefore, in this study, to solve this type of problem, in addition to solving explicit requirements through user-initiated query answering, an information recommendation is realized by the system. Thus, we develop a recommendation-type dialogue system that addresses the latent needs of prospective students and assists them in gathering information.[1] The information to be recommended (hereafter referred to as a "recommendation text") is provided in query form based on the preference analysis results using past and real-time query history and recommended information from the university, considering the personal attributes of the users. Information recommendation by the system is effective when the information that the user should grasp is unclear or when information that the user is vaguely interested in is efficiently collected.

2. Overview of Dialogue System to Be Developed

2.1 Positioning of Related Studies and Present Research

In natural language processing and artificial intelligence (AI), the use of dialogue systems as systems technology, practice reports, and design research is increasing[2]. Examples include voice car navigation systems, voice assistants, AI speakers, chat dialogue systems, and dialogue robots, which provide various functions, such as destination settings, music searches, open domain chat systems, and closed query answering.

Dialogue systems are roughly divided into task- and non-task-oriented types. Task-oriented systems have the explicit goal of solving a specific task. For instance, they are used for flight information[3] and bus operation[4] guidance. However, in this study, in addition to actively requested information, it is also assumed that the users collect information they are vaguely interested in. Such related studies solve specific tasks and differ in terms of their application.

Non-task-oriented types do not have a specific goal and are aimed at chatting with users. For example, studies that improve the accuracy of chat responses include response generation using knowledge graphs[5], response selection based on comfort level estimation using dialogue history and user evaluation[6], and ranking learning[7]. Using this method, we can select responses that reflect user satisfaction. However, because these studies do not support gathering user information, the difficulty in gathering such information cannot be solved.

Studies have been conducted to support the gathering of user information through interactive information exchanges while both the system and the user take the initiative during a dialogue. For example, a study provided related information that matched the intention of the user from the immediately preceding response text through predicate argument parsing[8]. In addition, another study focused on named entities, providing useful information to users in query form from candidates for the previous query answer[9]. In these studies, responses were generated from previous system responses and response candidates, which are useful when prompting users to collect information on specific topics. However, this study deals with a wide range of information, such as the number of students and research content, which is irrelevant but can be helpful for those who wish to pursue higher education. Therefore, we must support the information gathering of those who wish to pursue higher education in various areas without being limited to a specific topic.

2.2 Configuration of Proposed System

In this study, considering the personal attributes of

[1] In this study, we significantly improved the design model from a previous study[1] and developed a new recommendation-type dialogue system. We then evaluated the presence or absence of the recommendation function and an introduction evaluation of public relations for those who wish to pursue higher education at our university.

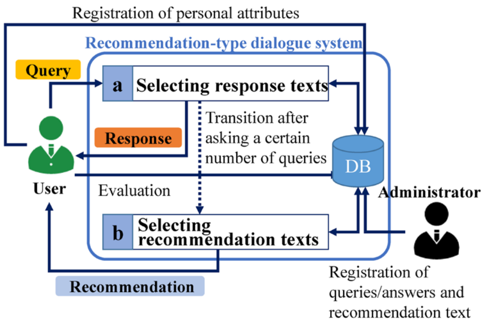

Figure 2. Overall configuration of the proposed system

U1	Are there student dormitories ?
S1	There are three dormitories.
U2	Tell me transportation.
S2	Mostly bicycles.
<<< Providing recommendation texts on other screen area >>>	
U3	How much is the apartment fee ?
S3	From 30,000 yen to 50,000 yen.
U4	How many students live alone ?
S4	About 60% students live alone.
<<< Providing~ >>>	
R1	How much is the apartment fee ? Do tsunamis occur ? How is the transportation?
R2	Tell me events in Tokushima. What is Awa-dance ? What is the famous foods ?

Figure 3. Example dialogue of the proposed system

users, potential user requests are evoked, and information gathering is supported by recommending a wide range of information that meets user preferences and recommended information from the university during a dialogue. The recommendation information is described in Section 3.2.

Fig. 2 shows the overall configuration of the dialogue system developed in this study. The DB records, among other information, the user-registered personal attributes, labeled queries/answers in the query categories registered by the administrators, recommended information, and query history. Personal attributes are described in Section 3.3, and the query fields are described in Section 3.4. The system manages dialogues using frames. Personal attributes, query text, query field, and query target (faculty/department) were used as the frame attributes.

For the normal queries and answers shown in Fig. 2(a), the personal attributes are acquired in advance, after which the morphological analyzer MeCab extracts words, such as nouns and exclamations, from the user query and determines the query target if necessary. The dictionary used for morphological analysis uses both the NEologd dictionary and a user dictionary created to determine words specific to the public relations of a university (e.g., common university entrance tests and APs). Personal attributes and query texts are recorded in the frame, and the query target is "Which faculty/department?" In addition, if there is missing information, the user is confirmed. After the frame is filled, the response texts are selected through text classification[2] using a distributed word representation. Distributed representation was obtained using a model created by Wikipedia. We prepared learning data from registered queries/answers and the dialogue history of the user and created a text classifier using fastText. In addition, to increase the correct answer rate, the query texts were word-converted prior to classification using a manually generated synonym/synonym dictionary. For example, "*Tokudai*" was converted into "Tokushima University" to resolve variations in spelling.

Next, as shown in Fig. 2(b), each time a user asks a certain number of queries, the preference analysis results and recommendation information are used to select a recommendation text that considers the user's personal attributes. Recommendation texts are provided in a query format using the query texts in the DB. The information recommendations are described in Section 3.

Fig. 3 shows an example of the dialogue used in this system, where U indicates the user's query, S is the system response, and R is the recommendation by the system. A wide range of information is provided by periodically sending recommendation text according to the content of the user query. In addition, to avoid disturbing the dialogue, a screen area for providing recommendations was created separately from the screen area where the query text of the user and the response text were displayed.

[2] To prepare the training data necessary for creating the classifier, at the beginning of the system introduction, the query texts in the DB that are most similar to the query text of the user are searched (cosine similarity), and the response text linked to that query text was selected.

3. Recommendation Method Using Preference Analysis Results and Recommended Information

3.1 User Preference Analysis

Recommendation texts are only used when they are of interest to users. A method for providing ranked information is available, which provides information that many users may be interested in based on the number of queries historically made. In this study, the information requested by users is diverse, and even a few queries may be helpful for some individuals. Therefore, it is necessary to provide recommendations that reflect the interests of each user.

It is highly possible that a certain user will also be interested in another user's query with similar tastes. For example, a user who is interested in AI might ask "At what kind of company did a student of AI find a job?" Such content would be helpful for other users interested in AI. Such information recommendations can promote awareness and the discovery of reference information.

Therefore, in this study, we use collaborative filtering[10] to select users with similar preferences from the accumulated preference information and provide information recommendations based on their preference patterns. This recommendation method is used in various situations, including product recommendations for online stores and article recommendations for internet news. This study used a user-to-user memory-based method, which is a representative collaborative filtering approach. This method calculates the similarity between a recommended user and other users and obtains the predicted evaluation value of a query text as a recommendation candidate based on the evaluation of similar users.

The similarity between users is calculated using the evaluation value given to query text i in the real-time query history of the recommended user and the past query histories of other users. The evaluation value is determined by whether a query exists and the user's evaluation: a value of 2 indicates a query and a high evaluation exists, 1 indicates a query and no evaluation exists, and 0 indicates a query and a low evaluation or no query or evaluation exists. The user evaluation method is described in section 3.5. Equation (1) shows the method used for calculating the degree of similarity between users. Let u be the recommended user, u' be another user, $r_{u,i}$ be the evaluation value of u for query i, and n be a set of query texts. The similarity $s(u, u')$ between u and u' is calculated using cosine similarity.

$$s(u,u') = \frac{\sum_{i \in n} r_{u,i} r_{u',i}}{\sqrt{\sum_{i \in n} r_{u,i}^2} \sqrt{\sum_{i \in n} r_{u',i}^2}} \tag{1}$$

$$E_{u,i} = \frac{\sum_{u' \in U} s(u,u')(r_{u',i})}{\sum_{u' \in U} s(u,u')} \tag{2}$$

Next, the predicted evaluation value of the query text is calculated using Eq. (2). Let U be the set of the top 30 other users, similar to the recommended user u. In a preliminary survey, no significant change occurred in the content of the recommendation texts, even when other users were limited to the top 30 individuals. They predicted that the evaluation value $E_{u,i}$ for query text i of recommended user u is calculated using the sum of the evaluation values of other users weighted by their similarity with the recommended user and the weighted average of the total similarity. In conventional collaborative filtering, query text with a high prediction evaluation is recommended.

3.2 Calculation of Final Prediction Evaluation Value Using Recommended Information

The order in which recommendation texts are provided depends on the evaluation value. Therefore, if the number of queries regarding the information that the university wants to ask prospective students is small and the user evaluation is low, the recommendation text might not be provided to the user. For example, the AP specifically shows an "image of a person the university is looking for" and the "priority of the evaluation items in the applicant selection method." However, there is a high possibility that the number of queries regarding the AP is small because recognition is not high. In fact, only a few students considered AP when taking an exam[11].

Therefore, we implemented a mechanism that preferentially recommends specific query texts by intentionally intervening in the evaluation of the university (e.g., the administrator of the proposed system) without depending on the evaluation value. Specifically, an administrator selects a query text in advance as the recommended information and then increases the predicted evaluation value of the query text, as shown in Eq. (3). Here, $N_{u,i}$ ($0 \leq N_{u,i} \leq 1$) is the final prediction evaluation value for query text i of recommended user u, D is the set of recommended recommendations (including the provided recommendation information), S is the set of recommended information, and E_{min} and $E_{\max}$ are the minimum and maximum values of $E_{u,i}$. In addition, $N_{u,i}$ is 0 for recommended texts that have already been provided, 1 for recommended information, and a normalized predicted evaluation value (maximum value of 1, minimum value of 0) for other cases. The final predicted evaluation value $N_{u,i}$ of the recommended information is the maximum value within the value range ($0 \leq N_{u,i} \leq 1$), and is therefore preferentially recommended.

$$N_{u,i} = \begin{cases} 0 & if\ i \in D \\ 1 & else if\ i \in S \\ \dfrac{E_{u,i} - E_{min}}{E_{max} - E_{min}} & otherwise \end{cases} \tag{3}$$

3.3 Information Recommendation Considering Personal Attributes

Depending on the time of year, users may not need information, such as events and entrance examinations, that are held indefinitely or for a specified period. For example, a college counseling session is held in January yearly. Providing this information to third-year high school students from February onwards is not helpful. In addition, information on Tokushima Prefecture and female students may be unnecessary, depending on the place of birth and gender. Therefore, to provide information in which the user may be more interested, we conduct an information recommendation that considers personal attributes.

Personal attributes include status (e.g., first through the third year of high school, graduate, or guardian), place of birth, and gender. We then set the recommended condition of the query text based on these attribute values and the provisioning time. Table 1 shows an example of the non-recommendation conditions. All conditions were AND conditions. For example, the results of a general entrance examination (scheduled for the first semester) will be announced in early March; thus, information regarding the schedule for the first semester will not be recommended for mid-March to third-year high school students and previously graduated students. The validity of the provisioning period will be on a yearly basis.

Table 1. An example of non-recommendation conditions

Query text (simplified)	Condition
Open Campus	Status (3rd year high school or already graduated)
	Season (after September)
Special Entrance Examination (Recommendation I)	Status (3rd year high school student or already graduated)
	Season (after late December)
General Entrance Examination (1st semester schedule)	Status (3rd year high school or already graduated)
	Season (after mid-March)
Introduction to Tokushima Prefecture	Birthplace (Tokushima Prefecture)
About Science Girls	Gender (male)

3.4 Recommendation Text Provisioning Method

It is desirable to increase the recommendation frequency to encourage users to collect more information. Previous studies[8], [9] provided information by detecting silence from a user for a certain period of time. However, continuously making information recommendations during a dialogue could increase opportunities for the user to show interest and spread a topic. Therefore, this system provides three recommendation texts each time three queries are made. The number of queries and recommendations was updated with reference to user evaluations after introducing the system.

At the counseling sessions held thus far, the queries of the prospective students cover various fields, such as entrance exam topics and the living environment. In addition, queries are grouped by field, including queries about life and entrance examinations. Such tendencies are the same in the dialogue system. Therefore, it is highly likely that the user will feel a sense of incongruity when a recommendation text is provided for a field that differs from that being investigated. For example, it is unnatural to provide a recommendation text asking “How many students live alone?” after being asked a query “Please tell me about the subjects for the general entrance examination (first semester) of the faculty of science and technology.”

Therefore, in this system, the query fields are classified into "features, entrance exams, life, and others," and recommendation texts in the same fields as the most recent queries of the user are provided. Specifically, among the three most recent queries, the final predicted evaluation value of the query text labeled with the most frequently asked query category is calculated. Here, "others" corresponds to greetings and queries to Tokupon (see Section 3.5 for details) and are not recommended.

3.5 Development of Dialogue System

Dialogue systems equipped with interfaces, such as humanoid robots and anthropomorphic agents, have been developed to allow users to feel familiar with the system[12]. Therefore, in this system, the mascot character "Tokupon" of Tokushima University is used as a responder. This system is called "Tokupon talk."

For the site configuration, a service introduction page is first prepared, and after selecting personal attributes and the number of usages, users transition to this system. Fig. 4 shows the reference screen for this system. The screen contains a recommendation list, query input form, and dialogue screen from the top. Because Tokushima Prefecture is the birthplace of the indigo, we used indigo as the background color. Recommendation, user query, and response texts are displayed by updating a partial screen area through asynchronous communication. The recommendation list displays the recommendation texts and response buttons. By pressing such a button, the response text corresponding to the recommendation text is displayed on the dialogue screen.

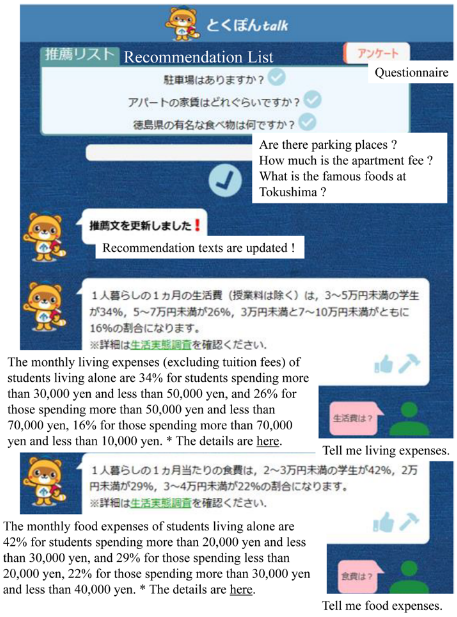

Figure 4. Tokupon talk reference screen

The dialogue screen mainly displays the user's query and response texts. To reduce the inorganic nature of the system, the time from the query to the display of the response text was irregular (0.05 to 0.5 s). In addition, to grasp the appropriateness of the response content and incorrect responses, a high/low evaluation button was added at the end of the response text.

The user is notified that the text has been updated during the dialogue to note whether the recommendation information has been updated. However, the user may not be able to evaluate the importance of the recommended text. Therefore, when providing a specific recommendation text such as "The AP contains important evaluation items for the entrance examination," the reason for the recommendation is noted. Additionally, a questionnaire button was installed at the top of the screen for user evaluation.

To respond accurately to user queries, a sufficient number of query/answer texts must be registered. However, there is information that can only be quickly grasped by the person in charge, such as irregularly held events and changes in public information. Therefore, we developed a management site with functions for registering and modifying queries and answers, and for viewing recommended information and query history. Fig. 5 shows a reference screen. A decentralized management system is arranged by hierarchically organizing multiple registration personnel under the administrator. Variations in the registered content quality may affect the correct answer and usage rates of the recommendation function. Therefore, with the objective of unifying the registration standards, this issue was resolved by providing the person in charge of registration the right to view all registered queries and answers and by introducing a mechanism that allows the administrator to verify the updated content at any time.

The software shown in Table 2 was used to develop Tokupon talk. A web system was developed using PHP, jQuery, and other tools in the virtual environment of the VMware ESXi. Assuming that the usage rates of smartphones and personal computers are high,

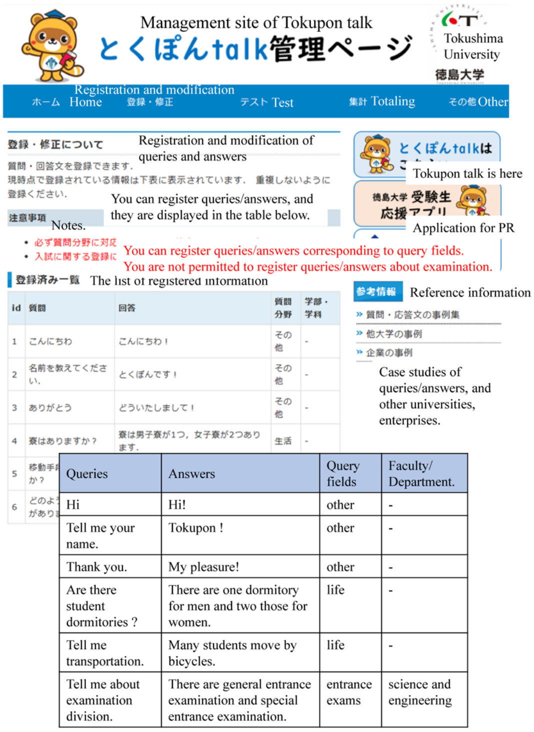

Figure 5. Reference screen of management site

Table 2. Software list

Type	Software
Virtualization Software	VMware ESXi 6.7.0
OS	CentOS 7.4 64bit
Http server	Apache 2.4.6
Database	MariaDB 5.5.52
Programming language, etc.	PHP 5.4.16, Python 3.6.8, jQuery 3.5.1, HTML
Other	fastText 0.9.2, MeCab, shibboleth 3.0.4

we designed screens for several types of terminals using a responsive web design. For the management site, we used the integrated authentication system of our university (Shibboleth authentication)[13] to implement the authentication and authorization functions based on our university account.

4. System Evaluation

We evaluated the presence or absence of a recommendation function and an introduction evaluation targeting public relations for those who wished to enroll at the university. The evaluation was conducted in May 2020 for the former and from April 4, 2019, to March 31, 2020, for the latter.

4.1 Dataset

We registered 458 query/answer texts in a dataset.[3] By referring to the university guidebook, selection guidelines, websites, and other areas of information, we manually registered information that would attract the attention of prospective students, such as the number of applicants and study-abroad programs. As recommended information, we selected two items, "What is the AP?" for describing the desired personality and important evaluation items in the entrance examination, and because introductions to the prefecture are popular information at educational consultation meetings, "What kind of place is Tokushima Prefecture?" Both are notifications of the reason for a recommendation, and only the latter is subject to the non-recommended condition of the hometown (Tokushima Prefecture).

4.2 Evaluation of the Presence or Absence of Recommendation Functions

To investigate the influence of the presence or absence of a recommendation function, we conducted a comparative evaluation using a system that does not present recommendation text (A) or the proposed system (B). The screen designs, functions, and the number of registered query/response texts were the same for both systems, except for the presence or absence of the presentation of recommendation texts. The number of queries was used as an evaluation index for the system. A system with numerous cases can be considered to have collected a large amount of information. The subjects were 24 third-year high school students from prefectures other than Tokushima Prefecture who wished to go to university. The participants were divided into two equal groups using random sampling. To avoid the influence

[3] Additional queries and answers were registered through an operation, and this number is the number of registrations in the final month of the evaluation (March 2020). We evaluated the presence or absence of a recommendation function based on this number.

of the evaluation order, one group was evaluated in the order of A to B, and the other group was evaluated in the order of B to A. For the evaluation task, assuming they were gathering information for the first time, we asked the students to complete an approximately 30-min evaluation. As a flow of evaluation, we asked the participants to prepare query texts in advance and input these query texts for the first and second times. Here, if a new query appeared, we could make an additional query only for the system under evaluation (A or B). For example, if A is evaluated, additional queries can only be asked to A. At least three queries were required to obtain the information recommendations. In addition, 580 users and 1,647 query histories were registered, and this evaluation was conducted such that a preference analysis could be sufficiently conducted.

As a result of this evaluation, the average number of queries (standard deviation equal to that below) was 5.08 (4.36) for A and 9.5 (9.15) for B. A total of 58.3% of the subjects used the recommendation text, and the average number of uses per person was 5.71 (9.36). A Wilcoxon signed-rank test ($p<0.01$) showed a significant probability of 0.003, indicating a significant difference in the average number of queries between A and B. Therefore, more information was collected when using the recommendation function.

In contrast, the average usage rate of the recommendation texts (percentage used in all recommendation texts provided) was 45%, whereas the average usage rate of the recommendation information was 20%. The reasons for the low average usage rate were limited interest in the entrance examination because the evaluation period is in May and there was no interest in Tokushima Prefecture because the subjects were not considering enrolling at a local university. However, none of the subjects asked for the same query as the recommended information. In particular, because the recognition of AP is low, specific information can be obtained by providing it as recommended information.

Next, we conducted a questionnaire survey to investigate the effectiveness of the recommendation function and update frequency. The following four queries were asked and evaluated on a 5-point scale (1= disagree to 5=agree).

(1) Compared to A, did B collect more reference information when deciding where to go?
(2) Compared to A, B collected more reference information when it was unclear what queries to ask.
(3) Was it possible to collect reference information that could not be obtained without recommendation (B only)?
(4) Are the frequencies of recommendation updates and number of displays appropriate (B only)?

Items (1) and (2) are relative evaluations of B compared with A, and (3) and (4) are absolute evaluations of B. For (2), considering subjects who did not fit the assumed situation, only those who did were evaluated. The average values for each query were (1) 3.83 (0.9), (2) 4.16 (0.59), (3) 4 (0.91), (4) 4.26 (0.94), and 19 respondents (2). Comments from the subjects included, “there was helpful information in the recommendation information,” “at B, I was able to grasp a lot of information during a short period of time,” and “I was able to collect information that I could not ask through recommendations.” Based on the results of (2) and (3), the recommendation function evoked potential requests and encouraged awareness and discovery of information that would help choose further education. In addition, from the results of (4), the update frequency and the number of recommendations displayed were appropriate.

From the above results, the recommendation function could contribute to solving the challenge of gathering information for those who wish to pursue higher education and support information gathering in choosing higher education. However, there were people who did not use the recommendation function, possibly because reference information was not provided. In addition, there was a comment that “the recommendation function is not necessary if you have already decided what you want to look up.” In this evaluation, there were five subjects who did not correspond to (2), and for users who had clear information they wanted to collect, the recommendation function may have hindered their information collection. Therefore, the average value of (1) is considered lower than that of the other queries. To solve these problems, it would be effective to implement recommended manual updates and non-display functions.

4.3 Introduction Evaluation

Tokupon talk was published on AP Navi, a site for examinees provided by the university. This system remains open to the public even after the evaluation period.

First, the number of users (cumulative) during the

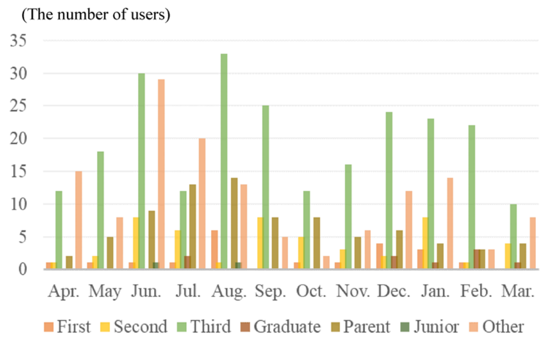

Figure 6. Number of users per month

period was 533, with the breakdown being 20 first-year high school students, 49 second-year high school students, 237 third-year high school students, nine graduates, 81 parents, two junior high school students, and younger, and 135 others. Of these, 66 users answered no queries. Fig. 6 shows the monthly number of users.

During the first half of the year, the number of users increased in June and August. In June, Tokupon Talk was introduced in a newspaper article, and the distribution of leaflets during counseling sessions for higher education is also considered a major factor for this increase. In addition, an open campus event of our university is held in August every year. Thus numerous queries were asked during the same month regarding this event, such as "When will it be held," "What kind of clothes are appropriate," and "Will participation or non-participation affect the entrance examinations?"

During the second half of the year, the number of users increased from December to January. The application period for the general and special entrance examinations (Recommendation II) held in 2019 was from late January to early February. Therefore, as a reference when choosing which school to attend, there were many queries from third-year high school students who were preparing to take their entrance exams, such as "What is the average score of successful applicants," "What are the characteristics of the faculty of science and technology," and "What is the percentage of people living alone?"

On the AP Navi exam site, the number of accesses increased by approximately 6-fold before and after the center test on January 18 and 19, 2020. This is due to the collection of information on realistic application candidates; however, the number of Tokupon talk users did not significantly increase. After the center test, there were numerous specific queries regarding the entrance examination, such as the entrance examination schedule and the method for announcing the results. However, because such information is published on the entrance examination site, AP Navi, and the university's website, Tokupon talk may not have been used as much.

Because an accurate evaluation becomes difficult, users with zero queries are excluded (e.g., the average number of queries and recommended texts used). The number of queries during the period was 1,508, of which 550 were regarding characteristics, 357 were on entrance exams, 421 focused on life, and 180 were about other subjects. The average number of queries per user was 3.83 (4.95).

After manually confirming all queries and answers, the correct answer rate was 74.8%. Although the correct answer rate in April was 35.7%, the rate increased over the course of the introduction period owing to the continual registrations of additional queries and response texts and the enrichment of synonyms and synonym dictionaries.

We provided 849 (290) recommendation texts (including those that were used), the breakdown of which was 336 (93) regarding the characteristics, 156 (35) on entrance exams, 240 (89) focusing on lifestyle, and 105 (67) regarding other topics. The average usage rate based on the query area was 27.7% for features, 22.4% for entrance exams, 37.0% for daily life exams, and 63.1% for others. The "other" category had the highest rate of usage, and in particular, many recommendation texts such as "What is the origin of your name," "What is your favorite food," and "What is the mark on your stomach?" were used for Tokupon. Although these are not useful information items for choosing a school, they have contributed to improving the familiarity with the system and the sustainability of the dialogue. In addition, because Tokupon is a mascot character of the university, the overall impression of the university may have improved.

Among all the users, 21.8% used recommendation texts, and the average number of uses per person was 3.37 (4.47). The average number of queries asked by nonusers of the recommendation texts was 2.30 (1.92), whereas the average number of queries by users increased to 9.31 (7.77). The Mann–Whitney U test ($p<0.01$) showed a significant probability of 9.35×10^{-37}, indicating a significant difference in the average number of queries with or without the use of recommendation texts. Examples of a user query (U) and the recommendation text (R) applied are in the following series: "(U) What is the student dormitory like?" "(R) Where are the apartments where the students live?" "(U) Please tell me about the nearest facility." "(R) What kind of circles (club activities) do you have?" "(U) What is the atmosphere like?" The users' information gathering was supported by expanding the topics and providing new topics through information recommendations.

The usage rate of the recommended information was 54.5% for "What is Tokushima Prefecture like?" and 28.6% for "What is the AP?" Both were higher than the average usage rate of the recommendation texts provided in the corresponding query area (the former being lifestyle and the latter being the entrance examination). Consideration of personal attributes and preferential provisioning of recommended information and the reason for the recommendation contributed to the improvement in the usage rate of recommended information. Recommended information is considered effective in providing user information that will serve as a reference for their choice of school.

Next, we describe the questionnaire results. The following five queries were asked and evaluated on a 5-point scale (1=I do not think so to 5=I think so).

(1) Did Tokupon talk help you achieve your goals (e.g., considering where to go to school)?
(2) Do you think that the addition of Tokupon talk to conventional public relations media will make information gathering easier?
(3) Was the update frequency and the number of recommendations appropriate?
(4) Was the content of the recommendation appropriate?
(5) Were you able to collect reference information that could not have been obtained without this recommendation?

There were 18 respondents: two second-year high school students, nine third-year high school students, one graduate student, four parents, and two others. The average values for each query were (1) 3.28 (1.49), (2) 4.56 (0.71), (3) 4.28 (0.75), (4) 4.38 (0.77), and (5) 3.94 (0.86). The evaluation of (1) for each user was polarized, and the evaluation was divided according to the frequency of queries that were answered incorrectly or could not be answered. Particularly at the beginning of the introduction, the evaluation was low because an insufficient number of queries and response texts were registered. Based on the results of (2)–(4), the effects of introducing Tokupon talk and the design of the recommendation function were evaluated. In addition, from the results of (5), it is considered that the users could collect more information by applying the recommendation function. One user commented, "The recommendation gave me a better understanding of student life in the department of my choice," suggesting that we were able to provide useful information.

Among all users, one had zero queries; however, the AP Navi examinee site used the reference screen of the Tokupon talk as a banner, and it is, therefore, possible that they accessed it out of curiosity. For such users, although it is difficult to analyze their preferences, we could encourage them to ask questions by making information recommendations earlier. Therefore, it is effective to provide popular queries as examples.

Many of the users applied the Tokupon talk multiple times, and some tended to ask the minimum necessary queries over the course of the day; for example, during the first access, only questions regarding the schedule of the open campus were asked; during the second access, only queries about the application method were asked; and during the third access, only questions on the venue were asked. Because Tokupon talk analyzes user preferences, making information recommendations that show the user's interest is difficult unless multiple queries are given concurrently. In addition, newly registered queries and answers may not be recommended because a query history has not yet been accumulated. This is called the cold-start problem[14], in which recommendations are not made until the evaluation data accumulate when the number of new users or items increases. The problem of new users may be solved temporarily using the query history of users with the same attribute information and who have made the same query. In addition, reinforcement learning defines the state action value and advances learning to maximize this value. By applying reinforcement learning, we could infer from attribute information and a small amount of query history information that can be helpful to users. The problem of new items might be solved by specifying the related query text and recommending the newly registered query text to the user who asks the question. In addition, by preferentially recommending the as-recommended information, the query history can accumulate.

In this evaluation, because the number of accumulated users and query histories was small at the beginning of the introduction, limited information that would be useful for users was recommended. In fact, among the 18 respondents to the questionnaire, the average values of the nine who answered earlier were (4) 4 (0.82) and (5) 3.67 (0.67), whereas the average values for the remaining nine subjects were (4) 4.78 (0.42) and (5) 4.22 (0.92). The average number of users who responded earlier was slightly lower. If helpful information is not recommended, the credibility of the users declines. Therefore, in the early stages of the introduction, until the number of users is accumulated, we must take measures such as presenting only information considered particularly important, such as university features and entrance exams, to avoid this problem.

The appropriate timing for a recommendation differs for each user. In this study, a recommendation is based on the number of queries; however, by considering the content and order of such queries, information recommendations can be made at the appropriate time for each user. Studies applying feedback during and after a dialogue to estimate levels of comfort and satisfaction have been conducted[6], [15]. In a future study, we must obtain more detailed feedback from users by referring to these approaches and considering the timing of the recommendation and a method for estimating appropriate recommendation content through a statistical analysis of the accumulated dialogue history.

The Tokupon talk presented three recommendation texts each time three queries were asked. Therefore, if four or more items of recommended information are registered, the information after the fourth item is not presented unless the user asks six or more questions. To present important information more reliably, we propose a mechanism that dynamically changes the presentation order of the recommended information according to priority settings and user queries. In addition, although recommendation texts in the same field as the most recent user query are provided, there is a possibility that the user's interest can be further aroused by providing information in a multifaceted manner regardless of the field. To support wide-ranging information gathering, we must evaluate the recommendations without limiting the field.

In this study, we evaluated the presence or absence of a recommendation function based on the number of queries and a questionnaire survey. However, we could not fully evaluate whether individual recommendation texts were useful for collecting information that would serve as a reference for choosing a school. Therefore, we must conduct a more detailed investigation of the evaluation of each recommendation text and the types of queries that tend to be developed after the recommendation text is used. In addition, to measure the effectiveness of public relations, the university conducted a survey of enrollees each year. One of the survey items was the reason for choosing the university, and the effectiveness of this system can be evaluated further by investigating the degree of contribution of the recommendation function when limited to users.

5. Conclusions

In this study, we developed a recommendation-type dialogue system that supports information gathering by arousing the latent demands of prospective students and targeting the public relations of a university. Specifically, based on the results of a preference analysis using past and real-time query histories and recommended information from universities, we implemented a function that recommends information that can be helpful to the user by considering their personal attributes in a query form. In the evaluation of the presence or absence of the recommendation function, the average number of queries (standard deviation) was 5.08 (4.36) without a recommendation, which increased to 9.5 (9.15) with a recommendation. In addition, the results of the questionnaire confirmed that the recommendation function evoked potential requests and that the awareness and discovery of information that would serve as a reference for deciding on further education were promoted. By introducing this system into public relations for students who wish to pursue higher education at our university, the average number of queries was 2.30 (1.92) for nonusers of recommendation texts, which increased to 9.31 (7.77) for users of such texts. Through these evaluations, we hope to resolve the difficulty in collecting information that occurs owing to the use of first-time entrance and passive examinations and support users in collecting information.

Future issues will include the following: methods for dealing with users who do not need or do not use the recommendation function, methods for promoting queries for users with zero questions, and a recommended method for users with few queries. Additional issues include an investigation into methods for estimating the recommended timing for each user, a redesigned method for presenting recommendation texts, and a more in-depth survey of the usefulness of choosing a destination. In addition, the query history accumulated in this system can be used for follow-up surveys after enrollment. Specifically, by collecting queries and profile information of prospective students, it can be used for enrollment management[16], realizing a university management cycle based on marketing methods. We could establish an effective student recruitment strategy by tracking pre-enrollment contacts with the university, examinations, admissions, and graduations. Therefore, we would like to consider a data analysis and tracking method by continuously introducing this system and a support method for further information gathering for those who wish to pursue higher education.

Acknowledgments

The authors would like to express their heartfelt gratitude to those in the Admissions Department and the Entrance Examination section of the Academic Affairs Department for their cooperation in registering the queries and answers for the Tokupon talk and preparing the introduction environment. Seki would also like to extend his gratitude to his wife for her help in designing the screen used in the proposed system. This research was supported by the JSPS Grant-in-Aid for Scientific Research (KAKENHI) 19K14317.

References

[1] Y. Seki and Y. Ueno, "A dialogue system recommending query sentences in consideration of user interest," *Proc. 2018 IEEE 7th Global Conf. Consumer Electron.*, pp. 428–430, 2018.

[2] H. Chen, X. Lju, D. Yin, and J. Tang, "A survey on dialogue systems: Recent advances and new frontier," *SIGKDD Explor.*, vol. 19, no. 2, pp. 25–35, 2017.

[3] V. Demberg and J. D. Moore, "Information presentation in spoken dialogue systems," *Proc. 11th Conf. Eur. Chapter Assoc. Comput. Linguistics*, 2006.

[4] L. Stanimirovic, Z. Cirovic, D. Donin, and M. Savic, "Serbian spoken dialogue system for bus travel information retrieval," *Eur. Signal Process. Conf.*, 2000.

[5] H. Zhou, T. Young, M. Huang, H. Zhao, J. Xu, and X. Zhu, "Commonsense knowledge aware conversation generation with graph attention," *Proc. Twenty-Seventh Int. Joint Conf. Artif. Intell.*, pp. 4623–4629, 2018.

[6] M. Mizukami, *et al.*, "Example based dialogue system based on satisfaction prediction," (in Japanese), *Trans. Jpn. Soc. Artif. Intell.*, vol. 31, no. 2, pp. 1–12, 2012.

[7] M. Inaba, N. Hirai, F. Toriumi, and K. Ishii, "Statistical response method for non-task-oriented dialogue agent," (in Japanese), *Trans. Fund. Electron. Commun. Compt. Sci.*, vol. 95, no. 6, pp. 1390–1400, 2012.

[8] K. Yoshino, S. Mori, and T. Kawahara, "Spoken dialogue system based on information extraction and presentation using similarity of predicate argument structures," (in Japanese), *J. Inf. Proc.*, vol. 52, no. 12, pp. 3386–3397, 2011.

[9] T. Misu, T. Kawahara, T. Shoji, and M. Minoh, "Speech-based interactive information guidance system using question-answering and information recommendation," (in Japanese), *Trans. Inf. Proc. Soc. Jpn.*, vol. 48, no. 12, pp. 3602–3611, 2007.

[10] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," *Proc. SIGIR Forum*, pp. 230–237, 1999.

[11] K. Nishimura, T. Inoue, and H. Nakamura, "What can be seen from the recognition situation of the admission policy: From the survey of prospective high school graduates in 2007 and 2016," (in Japanese), *Univ. Entr. Exam. Res. J.*, vol. 28, pp. 93–98, 2018.

[12] R. Nishimura, Y. Nishihara, R. Tsurumi, A. Lee, H. Saruwatari, and K. Shikano, "Operation of speech-oriented information system for real world research platform," (in Japanese), *IEICE Trans.*, vol. 87, no. 3, pp. 789–798, 2004.

[13] K. Matsuura, T. Ueta, and M. Sano, "Proposing user experience based on WEB-SSO with multiple backends of authentication/authorization engine," (in Japanese), *J. Acad. Compt. Netw.*, vol. 16, no. 16, pp. 138–145, 2012.

[14] A. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendations," *Proc. 25th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, pp. 11–15, 2002.

[15] S. Hara, N. Kitaoka, and K. Takeda, "Estimation method of user satisfaction using n-gram-based dialog history model for spoken dialog system," *Proc. of Int. Conf. Lang. Resour. Eval.*, 2010.

[16] S. Fukushima, "A case study of the comprehensive student information data analysis system: Enrollment management and institutional research at Yamagata University," (in Japanese), *J. Inf. Process. Manag.*, vol. 58, no. 1, pp. 2–11, 2015.

Yosuke Seki received his Ph.D. from Tokushima University in 2021. He is an associate professor at the Research Center for Higher Education at Tokushima University. He graduated from the Department of Science and Technology at Konan University. After a few years of working as a system engineer, he completed a Ph.D. program without a dissertation in systems innovation engineering at Tokushima University. His research interests include dialogue systems and the entrance examinations of universities. He is a member of JSET, JSAI, IPSJ, IEICE, and JACUAP.

Yoshihiko Ueno is a professor at the Research Center for Higher Education, Tokushima University. He graduated from the Meiji University. After working at a preparatory school, he completed a Master's degree at Oberlin University. His research interests include university entrance examinations. He is a member of the JACUAP.