Regular Paper

Baseless-Rumor Alert Bot to Promote Reliability of Information

Ryota Nishimura[1,a)]   Mai Miyabe Hirabayashi[2,b)]   Takashi Yoshino[3,c)]

Abstract: The dissemination of information by individuals, which has recently gained widespread attention, significantly affects multitudes of people regardless of the accuracy of the information. There is a vast amount of information available, making it difficult to identify baseless rumors, when the information is unverified. This study presents a bot called "Chillmo," which has been developed to alert users of such rumors. The proposed system aims to raise users' awareness of the reliability of information by actively alerting the users regarding the rumors receiving attention and providing the accurate information about such rumors through quizzes. Based on the experimental results, it was confirmed that Chillmo can make the users skeptical about rumors, promote the verification of authenticity, and raise their interest in the reliability of the information. In the future, we intend to conduct experiments with a larger number of people using the system to investigate trends by age and changes in the system usage over time.

Keywords: rumor, reliability of information, confirmation support, anti-diffusion, information literacy

1. Introduction

In recent years, social media has become a widespread and important tool for information dissemination and retrieval [1]. Social media allows any individual to easily post and share information, and send messages, containing valid information along with rumors [*1].

An unstable social situation, caused by natural disasters or other emergencies, generates a large amount of falsified information. With the emergence of the COVID-19 pandemic at the end of 2019, a number of rumors were spread in Japan, resulting in an adverse social impact that led to large-scale panics, such as buying toilet paper [2]. The former U.S. President, Donald Trump, made a statement to the effect that household disinfectants could be used to contain the coronavirus, which led to the U.S. government agencies issuing a warning regarding the use of cleaning agents to combat coronavirus [3].

Consequently, Twitter [*2], a microblogging service, has updated its guidelines and has taken measures such as deleting tweets with incorrect information regarding COVID-19 [4]. Various social networking services (SNS) such as Facebook [*3] and YouTube [*4] have taken similar measures [5] as well. However, it is difficult to distinguish between accurate and inaccurate information; moreover, excessive restrictions can threaten freedom of speech.

Rumors on these social networking sites spread to any receiver of the information as well as the users of the services. Even if the information is unreliable, it is accepted by some people and is then propagated further, and can lead to actions being taken based on such information, which is subsequently reported in the media. This implies that even if a rumor is sent out on Twitter, where the general age group of the users is relatively lower, it can be disseminated across people of different generations. Therefore, inaccurate information which was originally inaccurate can be considered true and acted upon, which can seriously affecting individuals and society. For example, as mentioned earlier, during the spread of COVID-19 in 2019, there was a shortage of goods due to hoarding. People believed that there would be a shortage of goods and ended up hoarding, which eventually resulted in the shortage of stock, despite the initial availability of sufficient stock. Similarly, there are a number of cases in which false information was considered accurate. To prevent such situations, it is necessary to create a system to make people aware of the uncertainty of daily information such as rumors, and to raise awareness of the reliability of the information. Therefore, we have developed a system called "Chillmo [*5]," which alerts users against such rumors. Chillmo is a chatbot that can respond in the natural language and can present the relevant rumors according to the user's query and alert the user of high-severity rumors. This paper presents the design principles and the functions of Chillmo, along with the experimental results and discussion.

---

1   Graduate School of Systems Engineering, Wakayama University, Wakayama 640–8510, Japan
2   Graduate School of Medicine and Faculty of Medicine, The University of Tokyo, Bunkyo, Tokyo 113–0033, Japan
3   Faculty of Systems Engineering, Wakayama University, Wakayama 640–8510, Japan
a)   nishimura.ryota@g.wakayama-u.jp
b)   mai.miyabe@gmail.com
c)   yoshino@wakayama-u.ac.jp

[*1]   In this study, the term rumor is defined as information that is not well founded and whose authenticity is unknown or questionable. It is irrespective of the malicious intent that exists in the process of generation.
[*2]   https://twitter.com
[*3]   https://facebook.com
[*4]   https://youtube.com
[*5]   We have named our developed system "Chillmo," which stands for "Check Rumor" and "Chill (Calm Down)."

2. Related Works

2.1 Studies Related to Information Diffusion

Kawakami stated that three factors, namely ambiguity, importance, and anxiety are strongly related to the transmission of rumors [8]. Allport et al. argued that the amount of rumor dissemination (R) is proportional to the product of importance (i) and ambiguity (a), and they proposed the formula $R \sim i \times a$ [9]. Knapp argued that the occurrence and transmission of rumors are related to social conditions that make many people feel insecure [10].

Wang et al. investigated the factors that people consider when sharing information [11]. In that study, interviews conducted among various American college students suggested that people share a news article not only because it is interesting to them and to the people that they share it with, but also because they find the content interesting. In addition, only approximately half of the participants mentioned that they would check for the validity of the shared information. However, almost all of the participants were unwilling to spread false information and expressed that they would retract the information that they shared if it was proven to be inaccurate.

Kakimoto et al. conducted a study to determine whether people take actions to confirm the accuracy of the information that is disseminated [12]. In their study, they conducted a questionnaire survey involving 108 Twitter users to verify whether Twitter users confirm the accuracy of the contents of the tweets before using the retweet function. They classified the tweets into four categories (tweets from real-world acquaintances, tweets from people known online, tweet from an acquaintance of an acquaintance, tweets from people who found it in a search) corresponding to their relationship with the sender and questions were asked, based on each of these categories. Consequently, for the three tweets other than the "Tweets from people who found it in a search," the percentage of people who validated the content of the tweet was less than 50%. For "Tweets from real-world acquaintances," the percentage was particularly low at 27%. This suggests that the information from people with whom we have a close relationship is less susceptible to suspicion. Furthermore, the percentage of teenage users who validate the content of a tweet is less than 20% for all four types of tweets, indicating that fewer users of this age group confirm the authenticity of information when compared to users in other age groups.

People do not always intentionally spread malicious rumors, but can spread them unknowingly, as they find them interesting or presume that they would be useful to others. However, this can often lead to the spread of dangerous information, such as incorrect medical knowledge, which can be life-threatening. In addition, only a few people validate the information that they have received. which can be attributed to the difficulty and the complexity of validation. This study presents a system that enables users to easily acquire information to identify rumors. The aim is to create a system that raises the awareness of the reliability of information in both sender and receiver.

2.2 Studies Related to Reliability Estimation of Information

Gupta et al. proposed a method to rank the reliability of tweets using supervised learning and a goodness-of-fit feedback method, which is one of the accuracy improvement methods in information retrieval [13]. They used features based on the content of the tweets, such as special characters, swear words, and emoticons, as well as features based on the user's information, such as the number of followers and the length of the username.

Qazvinian et al. proposed a method for the automatic reliability estimation of information on Twitter using a Bayesian classifier [14]. They used only the information available on Twitter, such as the content of the tweets, the user information, and unique information such as features for training. They experimented on more than 10,000 tweets that were manually labeled and showed that the proposed method achieves an MAP [*6] of 0.95 or higher.

Lim et al. proposed iFact, a system that automatically estimates the reliability of tweets [15]. They proposed a system that searches the internet based on the words used in the tweets and that automatically estimates the reliability of the tweets based on whether the results contain certain keywords (fake, doubt, etc.).

Because of the vast amount of information available on the internet, it is difficult to manually verify the authenticity of all the available information. Therefore, automatic estimation of the reliability of information is important to prevent the spread of rumors. However, because of the nature of misinformation, it is difficult to determine whether it is right or wrong. In addition, rumors do not necessarily spread exclusively through social networking sites and can be transmitted verbally from person to person. Therefore, it is necessary to not only examine the methods for determining the reliability of the information, but also develop a mechanism to raise awareness regarding the reliability of the information. We have developed a system that alerts users regarding rumors using rumor information collected by the "Rumor Cloud" [*7] [16].

2.3 Studies Related to Alerting Information Viewers

Ennals et al. proposed the Dispute Finder, a browser extension that alerts the user of controversial information on the internet [17]. Their system highlights the relevant words and phrases of the web pages that contain controversial content and allows the user to view the supporting and the opposing factors.

Saito et al. proposed a system to highlight information on a webpage, whose source is unknown [18]. The proposed system aims to help users browse the internet more carefully. The results of the experiment showed that the amount of time that the users spent searching and the number of web pages viewed, increased when highlighting was applied.

Schwartz et al. proposed a visualization method to help users determine the reliability of web pages [19]. Their method presents the percentage of web page viewers, who are experts in the page content. The experimental results show that the proposed method significantly improves the user's ability to determine the reliability of the search results, but it does not affect the reliability judgment on a page-by-page basis.

Various studies have been conducted to alert information viewers regarding the accuracy of the information, and an appropriate

---

[*6]   Mean Average Precision
[*7]   http://mednlp.jp/˜miyabe/rumorCloud/rumorlist.cgi

visualization can help the users determine the reliability of information. However, Gao et al. pointed out that the use of strong language or visualizations in presenting the reliability of information and opposing opinions can adversely affect the users' assumptions and result in the spread of false information [20].

Therefore, our system does not assert the authenticity of the information and only alerts the user of the inaccuracy of the information. This study provides the information to validate the authenticity of the information and encourages users to verify the information proactively.

3.   Survey Related to Information Sharing

A questionnaire survey was conducted on the frequency of the use of SNS and the actual state of information sharing to clarify the requirements for alerting people regarding rumors. Students at a Japanese university were asked to participate in this survey. The students were asked to request the cooperation of their families and acquaintances for the survey. The survey was conducted online via a questionnaire. **Table 1** shows the age and gender composition of the questionnaire respondents. Of the 178 respondents to the questionnaire, those in their 20s were the largest age group, accounting for 88% of the total. The male-to-female ratio was 4:6.

Figure 1 shows the results of the questionnaire regarding the frequency of LINE [*8], Twitter, and Facebook usage. Most of the respondents answered that they use LINE and Twitter more than twice a day, with 93% for LINE and 69% for Twitter. Conversely, 65% of the respondents answered that they did not use Facebook. These trends can be attributed to the fact that 88% of the respondents were in their 20s. According to the data from the Ministry of Internal Affairs and Communications (MIC), among the major SNS such as LINE, Twitter, and Facebook, LINE is the most widely used service among people of all ages in Japan. This system is intended to be used daily by people of various age groups. Considering the results of the questionnaire survey and the data from the MIC, the system is constructed as a chatbot that runs on LINE. Thus, we attempt to lower the threshold for users to use the system.

The results of the questionnaire on information sharing are shown in **Fig. 2**. In this questionnaire, "news" is defined as the information that is reported or published on TV or the internet, and a "rumor" is defined as the information whose validity is uncertain. Of the respondents, 87% said that they had shared "news" within a month, and 45% responded that they had shared "rumors" within a month. In both cases, the percentage of the respondents who answered that they had "shared" exceeded the percentage of those who answered that they had "not shared." The respondents who answered that they did not know whether they had shared "rumors" were 19%, which was considerably higher than in the case of the similar question regarding "news." These results suggest that there can be a certain number of people who are not aware of the validity of the information that they convey.

Figure 3 shows the partners with whom people shared "news" and "rumors;" 40% of the respondents shared "news" with "Par-
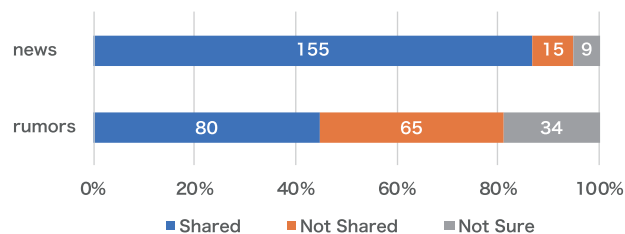
*8   https://line.me

Table 1   Age and gender composition of respondents.

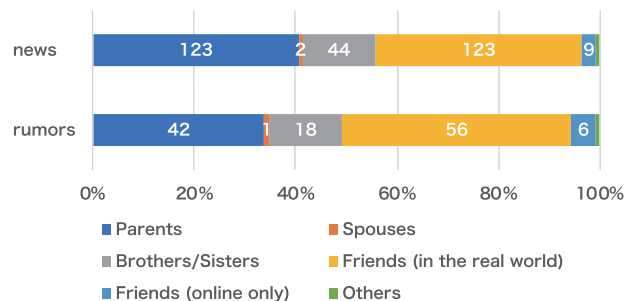| Age	Male	Female	Sum	Percent
10s	5	12	17	10%
20s	67	89	156	88%
30s	0	2	2	1%
40s	0	0	0	0%
50s	0	1	1	1%
SUM	72	106	178	100%

※ Percentages are calculated by rounding to the first decimal place.
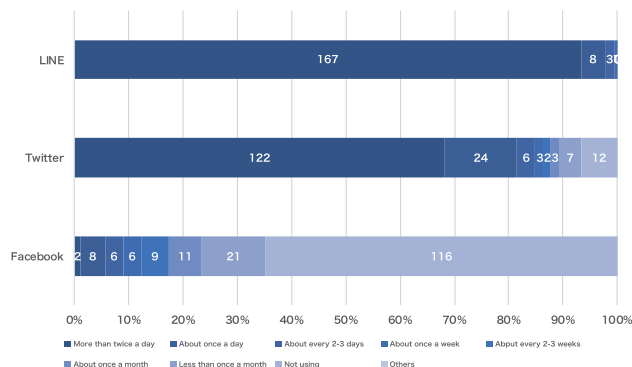
Fig. 1   Frequency of using SNS.

Fig. 2   Whether answers have shared news/rumors.

Fig. 3   With whom answers have shared news/rumors.

ents" and "Friends (in the real world)" respectively, while 34% and 46% shared "rumors" with "parents" and "friends in the real world," respectively. The results show that most of the respondents to the questionnaire shared the information they had acquired with a close circle of people around them. Essentially, there is a possibility that even people who do not use Twitter can receive rumors transmitted on Twitter through alternate communication channels. When information is shared among a close group of people, the validity of such information is rarely confirmed [12]. Therefore, it is necessary to alert not only the users of a social media service but also those who do not use that specific service.

The reasons for sharing "news" and "rumors" are shown in **Fig. 4**. In both cases, the fields, "The content was of interest to you" and "The content was likely to be of interest to the recipient," each accounted for approximately 40%. This is similar to
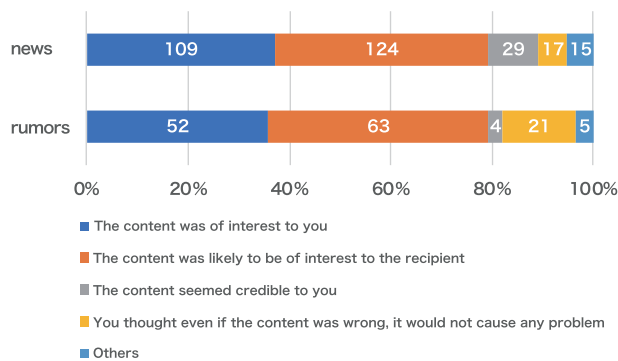
Fig. 4   Reason why answerers have shared news/rumors.

the results of a survey conducted on information sharing by Wang et al. [11]. Wang et al. concluded that people share information with others because they perceive a social need to be helpful. Even when people share information with good intentions, they can spread false rumors if they do not verify the authenticity of the information. Concerning the "rumors," 14% of the respondents assumed that "even if the content was wrong, it would not cause any problem," which is more than the similar response for "news." This suggests that there are times when people realize that the information they are about to share may be wrong, but still communicate it based on their judgment.

Based on the above results, this study proposes a system that provides a simple method to check for rumors by providing alerts of rumor information daily, and aims to build a system that raises the awareness of the danger of false information.

4.   Chillmo: Rumor Alert Bot

4.1   System Design Policy

As it is necessary to alert people of all ages, the system should be easily and routinely accessible. To sustain user continuity, it is necessary to reduce the burden on the users to acquire information regarding rumors and to verify their authenticity, as much as possible. Therefore, a chatbot system has been developed to allow users to text in a natural language to check for the rumors propagated on LINE, which is widely used in Japan. This system aims to raise awareness regarding rumors among the users and to enhance their interest in verifying the reliability of the information. The design policy of the system is summarized as follows:

(1) The system indicates the uncertainty of the information and encourages users to verify its authenticity.

Even if the system for detecting the rumors is implemented and operated, the users will have to determine the validity of the information by themselves, unless the detection accuracy reaches 100%. However, as mentioned in Section 2.3, if the system indicates the reliability of the information to a user in a definitive manner, there is a risk that the user conversely exacerbates their assumption. Therefore, in this study, information is presented to the users as being "possibly wrong," rather than as a definitive when alerting users to information. By raising the awareness of the uncertainty of the information among the users, this study aims to make users verify the information thereby encouraging them to make their own decisions regarding information dissemination.

(2) The system simplifies the process of verification process of information for the user.

The proposed system is built as a chatbot. The system extracts and presents rumors related to the user's input by using natural language processing techniques. Thus, users can easily verify the validity of the information that they have seen or heard by simply entering it into the system.

(3) The system encourages the users to use it daily.

Several websites can be used for fact-checking, including Factcheck.org [*9] and PolitiFact [*10]. However, as Wang et al. pointed out, there are very few people who perform fact-checking, and the existing mechanisms, such as fact-checking websites, are not routinely used. Therefore, we have developed a system that runs on LINE, which is widely used in Japan, so that the users can adapt the system comfortably in their daily lives. In addition, we consider methods to maintain the continuity of the usage of the system, such as appropriate frequency of notifications from the system.

4.2   System Configuration

The system configuration is shown in **Fig. 5**. "Chillmo" is a chatbot that runs on LINE. Therefore, users can start using LINE by adding Chillmo as a "friend" on their device. In addition, Chillmo can be used on any device that has LINE, including PCs, smartphones, and tablets.

When the user inputs text into the system, the "Confirm Rumor" (Fig. 5 (a)) function of the system is activated, and the system analyzes the input text and generates a response message. The input text is analyzed using Dialogflow [*11], a natural language processing platform provided by Google [*12], and the information such as the type of information required by the user is extracted as "Intention" from the input. The basic functions of chatbots are to greet people with a standard "Hello" and other similar salutations. If the input text includes "Confirm Rumor Keyword" such as "Is xxx true?," the "Acquire rumor information" (Fig. 5 (c)) function is activated. In this function, the system retrieves the information of the rumors that match or are related to the keywords entered by the user. The Alert rumor information (Fig. 5 (b)) function push-notifies the user of the rumors based on the information acquired in the Acquire rumor information (Fig. 5 (c)) function at regular intervals to alert the user.

The rumor information used in each function is acquired in conjunction with the Rumor Cloud, which has been in operation since 2012. Rumor Cloud periodically collects rumors from "Rumor Correcting Tweet" [*13] on Twitter. The information provided by Rumor Cloud includes the text of the rumor, "Rumor Correcting Tweet," "Number of Rumor Correcting Tweets," [*14]

[*9]   http://factcheck.org
[*10] https://www.politifact.com/
[*11] https://cloud.google.com/dialogflow
[*12] http://www.google.co.jp/
[*13] "Rumor Correcting Tweet" is a tweet from users on Twitter that denies or warns regarding rumors and is used by Rumor Cloud to make decisions when collecting rumors.
[*14] "Number of Rumor Correcting Tweets" is the number of "Rumor Correcting Tweets."
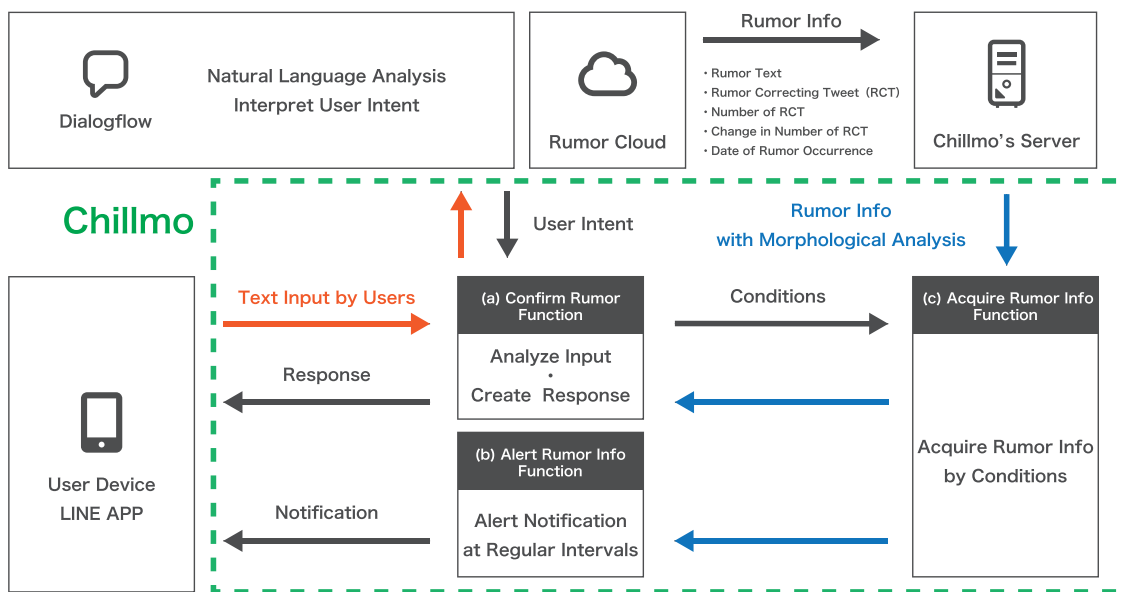
Fig. 5   System configuration.

and "Change in Number of Rumor Correcting Tweets." [*15] The "Rumor Information Acquisition Server" performs morphological analysis on the acquired rumor information and then passes this information along with the morphological analysis results to the Acquire rumor information (Fig. 5 (c)) function.

4.3   Basic Functions

The main screen of the system is shown in **Fig. 6** (a). The information of each rumor is presented in the form of a card. The user can check multiple pieces of information by sliding the card towards the left or the right, with a finger. A menu is placed at the bottom of the screen, and the user can call each of the functions described below by pressing the corresponding button.

Figure 6 (b) then shows how the bot responds to the messages input by the user. In this example, the system analyzes the user message, "I heard that bleach is effective against COVID-19" and presents multiple pieces of information based on related rumors. When the user clicks on the "Latest" button, shown in Fig. 6 (a), the system automatically sends the message "Tell me about latest rumors" and shows the user the rumors confirmed on that day.

When the user clicks on the "Spotlight" button, the system displays the rumors with the highest number of the "Number of Rumor Correcting Tweets" compared to the previous day in order, and when the user clicks on the "Ranking" button, the system displays the rumors with the highest number of "Number of Rumor Correcting Tweets" within a week in a ranking format. A user can input frequently used phrases with relatively low effort.

4.4   Linking to the Web Application

In addition to working as a chatbot, Chillmo works with a web application constructed with the LINE Front-end Framework (LIFF) [*16] to provide users with various functions. This section describes the functions available in web applications.
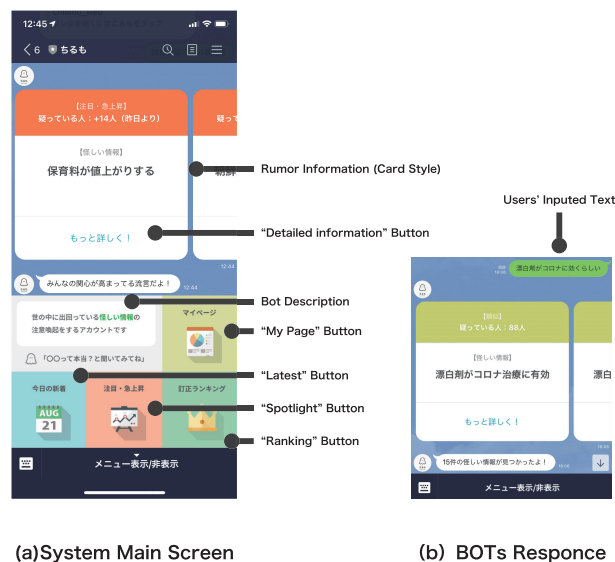

[*15]   "Change in Number of Rumor Correcting Tweets" is the number of "Rumor Correcting Tweets" compared to that of the previous day.
[*16]   The framework for building web applications that run on LINE.

(a) System Main Screen   (b) BOTs Responce

Fig. 6   Operation sample of basic functions.

4.4.1   Detailed Rumor Information

Figure 7 (a) shows the "Detailed Rumor Information" function. The user can view detailed rumor information by clicking on the "More" button at the bottom of the card (Fig. 7 (a)). This function presents the user with the "Rumor Occurrence Date," the "Rumor Correcting Tweet," the "Number of Rumor Correcting Tweets," and the "Change in Number of Rumor Correcting Tweets."

4.4.2   Number of Rumor Correcting Tweets Ranking

Figure 7 shows the "Number of Rumor Correcting Tweets Ranking" function. By clicking on the "Ranking" button (in the menu of Fig. 6 (a)), a user can view the "Number of Rumor Correcting Tweets Ranking," which shows the rumors with the highest number of "Number of Rumor Correcting Tweets" in a week in the ranking format. In this function, up to 200 words are displayed in the order of the number of rumor correcting tweets within a week. In addition, a user can view detailed rumor infor-

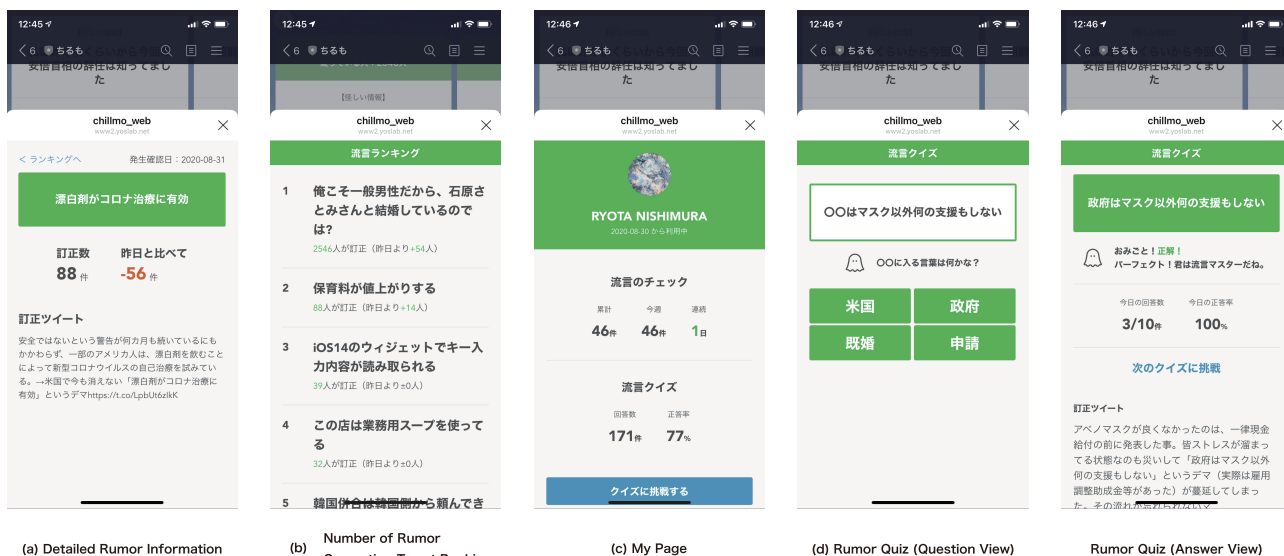| (a) Detailed Rumor Information | (b) Number of Rumor Correcting Tweet Ranking | (c) My Page | (d) Rumor Quiz (Question View) | Rumor Quiz (Answer View) |

Fig. 7　Operation sample of web applications working on "Chillmo."

mation by tapping on the displayed rumors (Fig. 7 (a)).

4.4.3 My Page

The "My Page" function aims to promote the continued usage of the system by providing a visualization function of the system usage status of the users and a quiz function about rumors. Figure 7 (c) shows the "My Page" function. A user can use the "My Page" function by clicking the "My Page" button (in the menu shown in Fig. 6 (a)). The "My Page" function presents basic information about the user, such as the date when the user started using the system and the number of "Detailed Rumor Information" that the user has viewed. The user can also use the quiz described below by clicking the "Challenge Rumor Quiz" button at the bottom of the screen.

4.4.4 Rumor Quiz

In the "Rumor Quiz" function, a user can learn about rumors with the help of a quiz. The purpose of this function is to promote the usage of the system by making it enjoyable. The "Rumor Quiz" function is shown in Fig. 7 (d) and the user can answer the quiz by clicking the "Challenge Rumor Quiz" button (Fig. 7 (c)). It is a quiz in which the user selects the content that applies to the foregrounded part of the text of rumors. The four choices displayed are automatically and randomly generated each time based on the results of the morphological analysis. After answering a quiz, the user can view the percentage of the correct answers and the "Rumor Correcting Tweet" used in the question.

5. Evaluation Experiment

5.1 Experimental Purpose

"Chillmo" aims to raise users' awareness of the reliability of information by alerting them to rumors. Therefore, the purpose of this experiment is to clarify the following.

Verification 1:　Whether "Chillmo" is effective in increasing users' interest in the reliability of the information.

As mentioned in Section 2.1 and Chapter 3, when people share information, they are rarely aware of the authenticity of that information. To prevent people from accepting such information and spreading rumors, it is necessary to raise awareness of the reliability of the information. Therefore, we verify whether the use of Chillmo is effective in raising users' interest in the reliability of the information.

Verification 2:　Whether "Chillmo" is effective in making users aware of rumors and encouraging them to voluntarily confirm the accuracy of the information.

As mentioned in Section 2.1, only a few people confirm the authenticity of information when it is shared. In addition, as discussed in Section 2.3, the use of strong language or visualizations in presenting the reliability of information can deepen the users' assumptions. It is not desirable for users to accept the content of the alerts on trust, because the accuracy of the content itself can change. Therefore, it is necessary to not only raise the user awareness of the alert information but also to trigger their spontaneous confirmation of its authenticity. In this study, we determine whether Chillmo can enable users to recognize rumors and promote spontaneous confirmation of the authenticity of the information.

Verification 3:　Whether "Chillmo" is a system that users can use continuously.

As discussed earlier, it is important to continuously verify the reliability of information when it is received or transmitted, because rumors occur and are spread daily. Therefore, it is desirable for users to continue using Chillmo and maintain a critical awareness of information. In this study, we examine whether Chillmo is a system that users can use continuously.

5.2 Experiment Procedure

In this experiment, two types of experimental conditions were used for comparison. The conditions of each experiment are as follows:

Condition A:　The "My Page" function is hidden, and a user cannot view the number of confirmed rumors or answer the "Rumor Quiz".

Condition B:　The "My Page" function is displayed, and a user can use all functions including "My Page".

Table 3   All questionnaire items.

ID	Question	Format
Q1	Throughout the period of using the system, your interest in the reliability of information has increased.	5-Point Likert Scale, Free Description of Reasons for Selection
Q2	Throughout the period of using the system, you have become more suspicious of the information you see.	5-Point Likert Scale, Free Description of Reasons for Selection
Q3	Throughout the period of using the system, you have increased your confirmation of the authenticity of information you see on a daily basis.	
Q4	Of the detailed rumor information, the "Number of Rumor Correcting Tweet" is important for confirming the information.	5-Point Likert Scale, Free Description of Reasons for Selection
Q5	Of the detailed rumor information, the "Change in Number of Rumor Correcting Tweet" is important for confirming the information.	
Q6	Of the detailed rumor information, the "Rumor Correcting Tweet" is important for confirming the information.	
Q7	Of the detailed rumor information, the "Rumor Occurrence Date" is important for confirming the information.	
Q8	You want to continue using the system after the experiment period is over	5-Point Likert Scale, Free Description of Reasons for Selection
Q9	"View Latest Rumor" function is necessary for the continuous use of the system.	5-Point Likert Scale, Free Description of Reasons for Selection
Q10	"View Spotlight Rumor" function is necessary for the continuous use of the system.	
Q11	"View Rumor for Ranking Format" function is necessary for the continuous use of the system.	
Q12	"Talk to Chillmo for Checking Rumor" function is necessary for the continuous use of the system.	
Q13	"Rumor Alert Notification" function is necessary for the continuous use of the system.	
Q14	"Detailed Rumor Information" function is necessary for the continuous use of the system.	
Q15*	"View the Number of Rumor Checks" function on "My page" is necessary for the continuous use of the system.	
Q16*	"View the Number of Rumor Quiz Answers" function on "My page" is necessary for the continuous use of the system.	
Q17*	"Rumor Quiz" function on "My page" is necessary for the continuous use of the system.	
Q18	With the notification, you are more likely to suspect the information you see on a daily basis.	5-Point Likert Scale, Free Description of Reasons for Selection
Q19	The notification was the trigger for using the system.	5-Point Likert Scale, Free Description of Reasons for Selection
Q20	How often you saw the notification.	1: Never, 2: 1 or 2, 3: 3 or 4, 4: 5 or 6, 5: All
Q21	How did you feel about the frequency of the notification.	1: Very Few, 2: Few, 3: Appropriate, 4: Many, 5: Too Many
Q22	How often would you like us to notify you?	Free Description
Q23*	Answering "Rumor Quiz" increased your interest in the rumors that appeared in the quiz.	5-Point Likert Scale, Free Description of Reasons for Selection
Q24	Describe what you liked about the system.	Free Description
Q25	If there are any additional features you would like to see, describe them.	

5-point Likert scale: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

※ Questions marked with asterisk were asked only to Group B because they were about "My Page" functions.

Table 2   Group distribution of participants.

Group	Male	Female	Sum
A	7	3	10
B	6	4	10
SUM	13	7	20

Group A: "My Page" Function is hidden and not available.
Group B: "My Page" Function is displayed and can be used.

As described in Section 4.4.3, the "My Page" function is intended to promote the usage of the system. Therefore, we examine whether users who have access to the "My Page" function respond more positively to the system than those who do not.

The experimental subjects were 20 undergraduate and graduate students from the Wakayama University. Among them, 13 were male and 7 were female (**Table 2**). We have divided the participants into two groups according to the conditions described above: "Group A," for which the "My Page" function was not available, and "Group B," for which the "My Page" function was available. The composition of each group is shown in Table 2. First, 7 out of 13 males and 3 out of 7 females were randomly selected and assigned to Group A. Then, the remaining 6 males and 4 females were assigned to Group B. Thus, we randomly assigned 10 participants each to groups A and B and the gender ratio distribution of the two groups was maintained as evenly distributed as possible.

Before starting the experiment, we asked the participants to add Chillmo as a friend on LINE and to indicate their gender using a form on the system. The participants were then asked to use the system for one week from September 1 to 7, 2020. During the experiment, no tasks were imposed on any of the participants, but they were prohibited from talking about the system with each other. The purpose of this constraint was to conduct a comparison experiment without informing the participants of the difference between the systems with the "My Page" function and those without the "My Page" function. During the experiment, the system sent push notifications once a day to alert participants of the rumors. After the experiment, the participants answered a questionnaire using a form on the system.

Table 3 lists all the questionnaire items. The participants from Group A were asked to answer 21 questions, excluding the questions about "My Page," and the participants from Group B were asked to answer 25 questions, including questions about "My Page." The content of the questionnaire includes the change in the participant's behavior owing to the use of the system and the evaluation of each system function. The survey was conducted anonymously. The unique ID of the participant was acquired on the system form, and the results of the survey were recorded in connection with the participant information such as gender, answered before the experiment term, and the browsing history of the rumor.

5.3   Experiment Results
5.3.1   System Usage

Table 4 shows the usage status of the "Talk to Chillmo" function. The average number of times a user confirmed a rumor by text input using this function was 0.24 times per day per user. The "Latest" (Fig. 4 (b)), the "Spotlight" (Fig. 4 (c)), and the "Ranking" (Fig. 4 (d)) buttons indicate the number of times that the "Latest," "Spotlight," and "Ranking" buttons in Fig. 6 (a) were

Table 4   Usage of "Talk to Chillmo" function.

Group		(a) User Input	(b) Latest	(c) Spotlight	(d) Ranking
A	Average	0.30	0.21	0.19	0.17
	Standard Deviation	0.37	0.20	0.16	0.14
B	Average	0.17	0.09	0.10	0.14
	Standard Deviation	0.26	0.16	0.18	0.25
SUM	Average	0.24	0.15	0.14	0.16
	Standard Deviation	0.21	0.13	0.13	0.17

※ The table numerics indicate the number of times a user uses the function per day. The unit is "times/person/day."

Table 5   Usage of functions.

Group		(a) Detail Info	(b) My Page	(c) Rumor Quiz
A	Average	1.26	-	-
	Standard Deviation	0.67	-	-
B	Average	1.60	1.05	0.59
	Standard Deviation	1.50	0.81	0.40
SUM	Average	1.43	1.05	0.59
	Standard Deviation	0.96	0.81	0.40

※ The table numerics indicate the number of times a user uses the function per day. The unit is "times/person/day."

Table 6   Questionnaire on the reliability of information.

Question		Group	Rating Distribution 1	2	3	4	5	Median	Mode
Q1	Your interest in the reliability of information has increased.	A	0	0	1	5	4	4	4
		B	0	0	2	7	1	4	4
		SUM	0	0	3	12	5	4	4

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

Table 7   Questionnaire on confirming the authenticity of information.

Question		Group	Rating Distribution 1	2	3	4	5	Median	Mode
Q2	You have become more suspicious of the information you see.	A	0	4	2	3	1	3	2
		B	0	2	2	4	2	4	4
		SUM	0	6	4	7	3	3.5	4
Q3	You've increased your confirmation of the authenticity of information you see on a daily basis.	A	0	2	4	4	0	3	3,4
		B	0	1	3	5	1	4	4
		SUM	0	3	7	9	1	3.5	4

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

clicked on, respectively. These functions have used an average of 0.15 times, 0.14 times, and 0.16 times for "Latest," "Spotlight," and "Ranking," respectively. Concerning the "Talk to Chillmo" functions, approximately 33% was based on the user's text input and approximately 67% was based on the buttons, indicating that the three buttons, which do not require any input, were frequently used.

Table 5 shows the usage of the "Detailed Rumor Information," "My Page," and "Rumor Quiz." "Detail" (Fig. 5 (a)) shows the number of times that the "Detailed Rumor Information" (Fig. 7 (a)) was viewed, and it was used 1.43 times per day per user. This is the number of times a user clicks on "More" (Fig. 7 (b)), and this can be used as one of the indicators of the user's interest in the rumors.

"My Page" (Fig. 5 (b)) shows the number of times a user used the "My Page" function (Fig. 7 (c)), which was used an average of 1.05 times per day per user. Group A is omitted from Table 5 (b) because the "My Page" function was not available for Group A. "Quiz" (Fig. 5 (c)) shows the number of times a user used the "Rumor Quiz" function (Fig. 7 (d) and (e)), which was an average of 0.59 times per day per user.

5.3.2   Interest in Information Reliability

Table 6 shows the results of question Q1: "Throughout the period of using the system, your interest in the reliability of information has increased." The Wilcoxon's rank sum test showed that the probability of significance between Groups A and B was 0.17, and there was no significant difference at the 5% level of significance. Therefore, we will now discuss the results of Q1 for both Groups A and B combined. The results for all the respondents were a median of 4 and a mode of 4. The following comments were recorded for those who responded with 5 and 4: "I realized that there is a lot of information out there, and that the content of that information is often based on the assumptions of the person who wrote it, so I became more interested in the relia-

bility of the information," "I've come to care about the source of my information," and "Since I was able to learn about rumors that I didn't know about through Chillmo, I became more interested in other information to see if it is reliable or not." It is likely that Chillmo alerts users of the uncertainty of the information due to the presence of rumors and raises their awareness of the reliability of the information. Conversely, one of the comments of those who responded with 3 was "I usually try to select reliable information." Thus, it is likely that those who mentioned that they had the habit of verifying information before using Chillmo did not provide a positive evaluation. This can be because the users did not realize a change in their awareness of the reliability of information owing to the use of Chillmo.

5.3.3   Confirming Information Reliability

Table 7 shows the results of the question Q2: "Throughout the period of using the system, you have become more suspicious of the information you see" and Q3: "Throughout the period of using the system, you have increased your confirmation of the authenticity of information you see on a daily basis." Wilcoxon's rank sum test was used to determine the probability of significance for Groups A and B. There was no significant difference at a 5% significance level for both Q2 and Q3. Therefore, we now discuss the results of Q2 and Q3 for both Groups A and B combined. The result for all the respondents was a median of 3.5 and a mode of 4 for both Q2 and Q3. The following comments were recorded for those who answered Q2 with 5 and 4: "I have grown more aware of whether or not the information I see is valid," "I have become more aware of rumors since I started checking for rumors." The following comments were recorded for those who answered Q3 with 5 and 4: "I tried to verify the authenticity of the false information provided by the system by myself," "Since there were some unexpected rumors, I have increased the number of times I question and confirm the authenticity of the information I get every day," and "I didn't have the means to confirm the authenticity of information before, but now I have more opportunities to confirm the authenticity of information with this system." It is likely that by viewing the information presented by Chillmo, users are more interested in verifying the reliability of the information, which leads them to confirm its authenticity. In addition, it was found that some people began to confirm the authenticity of the information that they saw and heard daily, not necessarily due to the information presented by Chillmo alone.

Conversely, some of the comments of those who responded with 3 and 2, were: "I've always taken the stance that I don't believe in anything suspicious, so I don't think the frequency of confirming authenticity has changed," "I tend to look at reliable information such as newspapers, etc. I did not confirm the authenticity of information I saw outside the system." These comments

were found to be common in Q2 and Q3. As in Q1, it is likely that for those who mentioned that they had the habit of questioning information before using Chillmo did not provide a positive evaluation.

Table 8 shows the results for Q4–Q7. These questions assessed the importance of each element of the "Detailed Rumor Information" function in confirming the information. The Wilcoxon's rank sum test was used to determine the probability of significance of Groups A and B for Q4–Q7. There was no significant difference between the two groups at a 5% significance level. Therefore, in the following sections, we will discuss the results of these questions for Groups A and B combined. Among the detailed rumor information, the median and mode values of Q4: "Number of Rumor Correcting Tweets," Q6: "number of tweets," and Q7: "date of occurrence" were 4 or 5, indicating that the information is required by users. In contrast, the median and frequency of Q5: "Change in Number of Rumor Correcting Tweets" were both 3, which is lower than the other information. The following comments were recorded for those who answered Q5 with 3 and 2: "I didn't notice much of a change in these numbers," "I don't think it would be helpful to present a fine numerical increase or decrease." Thus, in the present experiment, the user felt little change in "Change in Number of Correcting Tweet," and consequently, it was evaluated as not very necessary information. Therefore, in the future, it is necessary to identify "Change in Number of Correcting Tweet" to show only those items that have increased or decreased significantly.

5.3.4 Continuous Usage

The results of question Q8: "You want to continue using the system after the experiment period is over" are shown in **Table 9**. The Wilcoxon's rank sum test showed that the probability of significance between Groups A and B was 0.02, and there was a significant difference at a 5% level of significance. All of the users in Group B, who used "My Page," responded with 4 or 5, indicating that they are more positive about continuing the usage of the system than users in Group A. In the user comments of group B, the following were observed: "I found the system useful as a means of gaining knowledge," "Using the system helps us to remain aware of the authenticity of the information." It is likely that the easy detection of rumors through Chillmo led to a positive evaluation for continuous usage. Conversely, in Group A, those who responded with 3 and 2 commented: "I don't need to use the system because I have always checked the authenticity of news and other information by myself," "Because there wasn't much content of interest to me." Thus, it is likely that positive evaluations were not gained from those in the habit of acquiring information on their own or from those who could not acquire information of interest through Chillmo.

Table 10 shows the questionnaire on functions that users require for continuous use of the system. For questions Q9–Q14, Wilcoxon's rank sum test was used to determine the probability of significance for Groups A and B. There was no significant difference at the 5% significance level. Therefore, from now on, we will discuss the results of Groups A and B together in these questions. Q15–Q17 are related to the "My Page" function. Since the "My Page" function was available only to Group B, Q15–Q17

Table 8 Questionnaire on "Detailed Rumor Information" function.

Information		Group	Rating Distribution					Median	Mode
			1	2	3	4	5		
Q4	Number of Rumor Correcting Tweet	A	0	0	2	3	5	5	5
		B	0	0	1	5	4	4	4
		SUM	0	0	3	8	9	4	5
Q5	Change in Number of Rumor Correcting Tweet	A	0	2	5	3	0	3	3
		B	0	1	5	2	2	3	3
		SUM	0	3	10	5	2	3	3
Q6	Rumor Correcting Tweet	A	0	2	2	3	3	4	4.5
		B	0	1	2	5	2	4	4
		SUM	0	3	4	8	5	4	4
Q7	Date of Occurrence of a Rumor	A	1	0	1	3	5	4.5	5
		B	0	1	0	5	4	4	4
		SUM	1	1	1	8	9	4	5

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

Table 9 Questionnaire on continuous usege.

Question	Group	Rating Distribution					Median	Mode
		1	2	3	4	5		
Q8 — You want to continue using the system after the experiment period is over.	A	0	2	2	6	0	4	4
	B	0	0	0	8	2	4	4
	SUM	0	2	2	14	2	4	4

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

Table 10 Questionnaire on functions that users require for continuous use of the system.

Function		Group	Rating Distribution					Median	Mode
			1	2	3	4	5		
Q9	View Latest Rumor	A	0	0	0	2	8	5	5
		B	0	0	1	4	5	4.5	5
		SUM	0	0	1	6	13	5	5
Q10	View Spotlight Rumor	A	0	1	0	4	5	4.5	5
		B	0	0	0	5	5	4.5	4.5
		SUM	0	1	0	9	10	4.5	5
Q11	View Rumor for Ranking Format	A	0	0	2	4	4	4	4.5
		B	0	1	3	2	4	4	5
		SUM	0	1	5	6	8	4	5
Q12	Talk to Chillmo for Checking Rumor	A	0	1	2	5	2	4	4
		B	0	0	2	6	2	4	4
		SUM	0	1	4	11	4	4	4
Q13	Rumor Alert Notification	A	0	1	1	3	5	4.5	5
		B	0	1	0	5	4	4	4
		SUM	0	2	1	8	9	4	5
Q14	Detailed Rumor Information	A	0	0	1	6	3	4	4
		B	0	0	0	4	6	5	5
		SUM	0	0	1	10	9	4	4
Q15	View the Number of Rumor Checks	B	0	3	4	2	1	3	3
Q16	View the Number of Rumor Quiz Answers	B	0	1	5	1	3	3	3
Q17	Rumor Quiz	B	0	0	3	4	3	4	4

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

were asked only to the group. From the results of Q8, the participants of Group B were found to be more positive about the continued use of the system than the participants of Group A. However, for Q9–Q14, which are common to both Groups A and B, there were no significant differences in opinions regarding the necessary functions to continuous system usage between the groups. Therefore, the presence of the "My Page" function may have caused subjects to respond more positively to the continuous use of the system. In particular, the "Rumor Quiz" function received a rating of 4 in both median and the mode values, suggesting that it can improve the motivation of the subjects to continue using the system.

5.3.5 Rumor Alert Notification

Table 11 shows the results of the questionnaire on the effectiveness and frequency of "Rumor Alert Notification." For questions Q18–Q20, Wilcoxon's rank sum test was used to determine the probability of significance for Groups A and B. There was no significant difference at the 5% significance level. Therefore, from now on, we will discuss the results of Groups A and B together for these questions. In Q18: "With the notification, you are more likely to suspect the information you see on a daily basis," the median value was 3.5 and the mode was 4. In Q19: "The notification was the trigger for using the system," the median was

Table 11 Questionnaire on "Rumor Alert Notification".

	Question	Group	Rating Distribution 1	2	3	4	5	Median	Mode
Q18	With the notification, you are more likely to suspect the information you see on a daily basis.	A	0	4	3	0	3	3	3.5
		B	0	0	3	7	0	4	4
		SUM	0	4	6	7	3	3.5	4
Q19	The notification was the trigger for using the system.	A	0	1	0	1	8	5	5
		B	0	1	0	5	4	4	4
		SUM	0	2	1	8	9	4	5
Q20	How often you saw the notification.	A	0	0	6	1	3	3	3
		B	0	3	0	3	4	4	4
		SUM	0	3	6	4	7	4	5
Q21	How did you feel about the frequency of the notification.	A	0	0	7	3	0	3	3
		B	0	1	9	0	0	3	3
		SUM	0	1	16	3	0	3	3

※ The distribution of each evaluation is shown in Table 3.

4 and the mode was 5. The following comments were made as positive evaluations of these questions. "The daily notifications made me more aware of rumors on a daily basis," "Since I was notified about the rumors every day, I became more aware of them." In Q13: "Rumor Alert Notification function is necessary for the continuous use of the system." (Table 10) described in Section 5.3.4, the median value was 4 and the mode value was 5, indicating that the majority of the respondents rated the notification as a necessary function for continuous use. These results suggest that the notification function can serve as a trigger for the use of the system and leads to the continuous use of the system by users by encouraging them to suspect the information. From the results of Q20: "How often you saw the notification," more than half of the users saw at least five of the seven notifications. This indicates that this function was used frequently.

In contrast, the following negative comments were made regarding the notification function. "The contents of the notification were not of much interest to me," "There was not much of interest to me in the contents of the notification." For Q21: "How did you feel about the frequency of the notification," the Wilcoxon rank sum test was used to determine the probability of significance for Groups A and B, which was 0.04, indicating a significant difference at the 5% significance level. The number of users who rated the frequency of notification as "Appropriate" was significantly higher in Group B, while the number of users who rated it as "Too Many" was higher in Group A. For Q22: "How often would you like us to notify you?" (Table 3), the comments included "Once a day (same frequency as this experiment)," "Once every two days," and "A few times a week." Based on the result that Group B is more positive toward continuous use according to Q8 (Table 9), it is possible that providing the appropriate frequency of the notifications to each user can affect their willingness to continue using the system. Therefore, in the future, improvements such as "considering the contents of the notifications that are of interest to users" and "allowing users to select the frequency of the notifications" can enhance the effectiveness of the notifications.

6. Discussion

This section discusses the verification of items 1 to 3 based on the results of the experiments described in Chapter 5. The influence of the presence or the absence of "My Page" on each verification item is also discussed.

6.1 Verification 1: Whether "Chillmo" is effective in increasing users' interest in the reliability of information

The median and the mode values for Q1: "Throughout the period of using the system, your interest in the reliability of information has increased" (Table 6) were both 4. In addition, 17 out of 20 respondents answered 5 or 4, which implies that 85% of the experiment participants perceived that their interest in the reliability of information had increased. The following comments were made by the respondents who answered "Neither" in item 3. "Since I usually try to select reliable information, I have not become more concerned about the reliability of information than before," "The alert information also contained unreliable information." One of the characteristics of the users inferred from these comments is the possibility of them having a habit of questioning information even before using Chillmo. Essentially, these users can already be in a state in which they are highly concerned about the reliability of the information. Therefore, Chillmo can be particularly effective for people who are less likely to question information before using the system. Conversely, none of the respondents who answered 3 made negative comments about acquiring information through Chillmo; therefore, these people may also be positive about the use of the system itself.

In addition, the "Talk to Chillmo" function (Table 4) was used an average of 0.69 times per day per user when "User Input," "Latest," "Spotlight," and "Ranking" were combined. It is possible that the users are not only viewing the information notified by "Chillmo", but also actively accessing the information with an interest in rumors.

6.2 Verification 2: Whether "Chillmo" is effective in making users aware of rumors information and encouraging them to voluntarily confirm whether the information is correct

In Q3: "Throughout the period of using the system, you have increased your confirmation of the authenticity of the information you see on a daily basis." In (Table 7), the median value is 3.5, and the mode value is 4. Ten respondents, or half of the total respondents, answered 4 or 5. Conversely, 10 people rated the system as 2 or 3, indicating that they perceived the frequency of authenticity confirmation not to have increased after using the system when compared to before using the system. Therefore, we analyzed those who did not experience these changes. **Table 12** shows the ratings and comments of those who answered not positively to Q3. There was no significant difference between the groups in terms of the evaluation of Q3, but the percentage of those who made non-positive comments was slightly higher at six in Group A when compared to four in Group B. Based on the comments on the reason for the evaluation in Q3, it was discovered that several respondents were interested in the reliability of information and had a habit of verifying the information before using Chillmo. These people are marked with an asterisk (*) on the table. These people accounted for more than half (6) of the 10 respondents who answered negatively to Q3. These results indicate that Chillmo is effective in promoting validation among people who are not in the habit of verifying information, but it is not able to further strengthen this tendency

Table 12 Evaluation and comments of those who answered not-positively to Q3.

Participant	Group	Q3	Comment	Q8	Comment
1*	A	2	I tend to look at reliable information such as newspapers, etc. I did not confirm the authenticity of information I saw outside the system.	4	When the system notified me, I could check the rumors.
2	A	3	I found that there was a mixture of rumors in the information. But I did not doubt all the information.	4	I could use this system to help me remember the rumors.
3*	A	2	I usually check the source of information, so I did not feel that the use of the system changed the frequency of confirming the authenticity of information.	2	Even if I don't have the system, I usually examine the information I have.
4	A	2	I thought I'd just read the information that Chillmo would give me.	4	I found the system useful because I was able to use it without any stress and it provides me with information about the rumors every day.
5	A	3	I didn't bother to research the authenticity of the information myself.	4	I would like to continue using the system because it allows me to easily acquire rumor information and broaden my knowledge.
6*	A	2	Since the beginning of the experiment, I have changed my mindset to avoid sharing information with others without confirming its authenticity. However, during the experiment, there were no situations where I had to share information whose authenticity was unknown.	4	I would like to use it for a long time and see how "Number of Rumor Correcting Tweet Change Ranking" changes.
7*	B	2	Since I had always questioned the authenticity of information, there was no change in the frequency of confirming the authenticity.	4	Because by using the system, I can learn about topics that are being cared about in the world.
8*	B	3	I've always taken the stance that I don't believe in anything suspicious, so I don't think the frequency of confirming authenticity has changed.	4	I think it's useful if the system deals with rumors that are semi-universal in content.
9	B	2	I began to question the information, but I did not confirm its authenticity.	4	I found this system to be a useful tool to get information.
10*	B	3	Since I am usually careful about the authenticity of information, I neither increase nor decrease the frequency of verifying its authenticity.	4	Using the system helps us to remain aware of the authenticity of the information.

※ Participants who answered "I have a habit of questioning information even before using the system" in Q2 are marked with an asterisk (*).

among people who already have the habit of verifying information.

In addition, we discuss the evaluation of Q8: "You want to continue using the system after the experiment period is over" (Table 9) by those who answered negatively to Q3 (Table 12). Among these respondents, 9 out of 10 rated Q8 as 4. This indicates that most of the respondents, even those who have a habit of verifying information, responded positively to the continuous usage of the system. This can be confirmed by the comments of Participant 7: "Because by using the system, I can learn about topics that are being cared about in the world" and "Using the system helps us to remain aware of the authenticity of the information." Thus, Chillmo is effective as a means of acquiring information, even for those who have the habit of verifying information.

In addition, the user commented, "During this experiment, Rumor Alert Notifications sent by Chillmo often led me to use the system," and "I used the Detail Information function to confirm the authenticity of information that I found interesting in the system, but I did not search the internet for the authenticity of the information myself." There were several such comments referring to verifying the authenticity of the information by means other than Chillmo. Therefore, in addition to the information presented by Chillmo, it can be effective to assist users by presenting effective keywords for web search to promote the confirmation of the authenticity of the information.

6.3 Verification 3: Whether "Chillmo" is a system that users can use continuously

As mentioned in Section 5.3.4, the results indicate a significantly higher positive evaluation of the continued use of the system in Group B when compared to Group A. Regarding Q8: "You want to continue using the system after the experiment period is over" (Table 9), all the respondents in Group B answered 4 or 5. Considering Group A, both median and mode were rated 4, and 60% of the respondents were positive about its continued use. In addition, as described in Section 6.2, there is a possibility that Chillmo works effectively for users who did not experience any change in their awareness or behavior due to the system as a means to acquire information. Therefore, even if the proportion of users who are highly concerned about the reliability of information increases through the use of Chillmo in the future, there is a possibility that they will continue to use the system.

Figure 8 shows the time at which the detailed rumor information (Table 4 (a)) was viewed. Of the 200 times viewed, 109 times were between 8:00 pm and 9:00 pm, indicating that more than half of the total views were during this period. This is assumed to be due to the fact that Chillmo sent push notifications of the rumor information at 8:00 p.m. every day during the experiment. A user commented (Section 5.3.5) "I tried to verify the authenticity of the false information provided by the system by myself" and "When the system notifies me of rumors, I am more likely to be interested in it." This suggests that many of the respondents viewed the notification from Chillmo as a positive trigger for using the system. Based on these results, it is possible that push notifications can trigger and promote continuous usage.

As this experiment was conducted for one week, it is necessary to continue to analyze the usage trends of users by conducting experiments for a longer term.
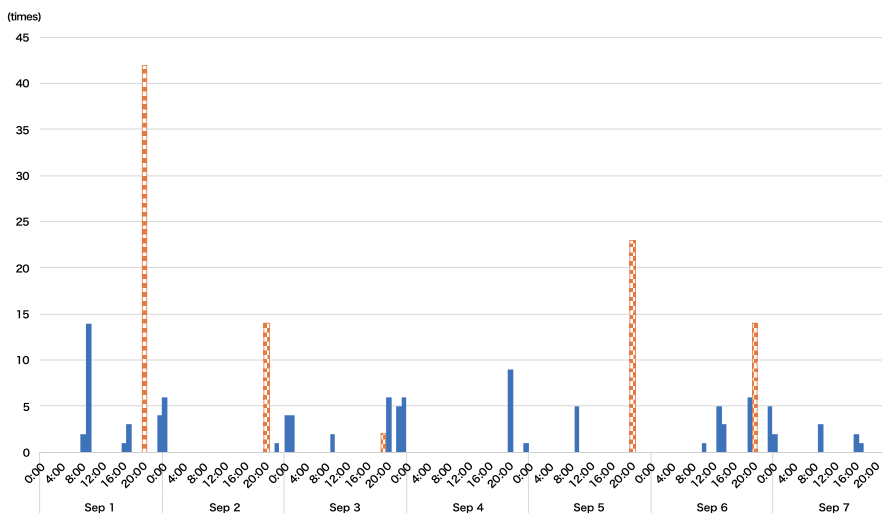
Fig. 8 The time and number of times the user viewed detailed rumor information.
※ The timing of the notification, 8:00 p.m., is indicated by an orange checkerboard pattern.

6.4 Influence of My Page Function

As mentioned in Section 5.3.4, in this experiment, there was a difference in Q8: "You want to continue using the system after the experiment period is over" depending on whether "My Page" was available or not, and Group B showed a significantly positive response regarding the continued usage of the system. In this section, we discuss the influence of the "My Page" function on the users.

The median and mode values for Q15: "'View the Number of Rumor Checks' function on 'My page' is necessary for the continuous use of the system" and Q16: "'View the Number of Rumor Quiz Answers' function on 'My page' is necessary for the continuous use of the system" were both 3 (Table 10). One respondent provided the reason for the evaluation as, "It motivated me to continue using the system." Conversely, more than half of the respondents commented that they did not pay much attention to the functions or did not use the functions. Therefore, it is unlikely that these functions would have sufficiently promoted the use of Chillmo. Q17: "'Rumor Quiz' function on 'My page' is necessary for the continuous use of the system." had a median and a mode frequency rating of 4 (Table 10). There were comments such as, "I thought this function would be a good motivation to continue using the system, since it would allow me to study rumors in a fun way," "Through the quiz, I could learn about rumors in genres you are not familiar with," and "Rumor Quiz function motivated me to use the system." which were stated as the reasons for the evaluation. Among the "My Page" functions, the "Rumor Quiz" function is an interactive function, which may have helped the users enjoy using the system. The "Rumor Quiz" function automatically generates questions from various quotations, regardless of the user's interests. As mentioned in the comments, the quiz often provided unknown information to the users, which led to discoveries. These results suggest that the "Rumor Quiz" function can provide users with an incentive to use the system. **Table 13** shows the results of the questionnaire on the effect of "Rumor Quiz" on the level of interest in rumors. Q23: "Answering "Rumor Quiz" increased your interest in the rumors that appeared in the quiz." resulted in a median value of 3.5 and a

Table 13 Questionnaire on the effect of rumor quiz on the level of interest in rumors.

Question	Rating Distribution 1 2 3 4 5	Median	Mode
Q23 Answering "Rumor Quiz" increased your interest in the rumors that appeared in the quiz.	0 1 4 2 3	3.5	3

Rating Distribution: 1: Strongly Disagree, 2: Disagree, 3: Neither, 4: Agree, 5: Strongly Agree

mode value of 3. Although the rating from those who did not use the "Rumor Quiz" function itself was low, the comments of those who answered 5 or 4 included the following. "I researched the information I learned through the quiz on my own," "Thinking about the correct answer to the quiz helped me to remember the rumor well." This suggests that quizzes can have the effect of making users more interested in rumors if they can be encouraged to answer them.

As these functions were not made known to the participants in this experiment, it was assumed that some of the participants were not aware of the existence of these functions in the first place. In addition, the number of consecutive quiz responses was found to be less than five questions per day. The mode of the number of times they can answer the quiz can be too high. Therefore, it is possible to encourage users to use the system by informing them of the functions of "My Page," and by reducing the number of possible answers per day to no more than five questions or reviewing the contents of the quiz questions.

7. Conclusion

This study presents a system that alerts people of rumors to increase their interest in the reliability of information. An evaluation experiment was conducted to verify the utility of the system. The contributions of this study are as follows:
(1) The proposed system can raise the user's interest and awareness regarding the reliability of the information. It has been demonstrated that the proposed system has the potential to promote the behavior of spontaneous confirmation of information, and the design policy of the system is confirmed to

be acceptable.

(2) There was no difference in the effect of the system function on the promotion of the behavior of confirming the authenticity of information. It was suggested that the information literacy of the individual is a factor that plays a major role in confirming the authenticity of the information, and considering the level of information literacy of the users can lead to the promotion of appropriate authenticity confirmation behavior.

(3) It was confirmed that the alerting of rumors by the system is effective in raising the awareness of the reliability of information among people who do not generally verify information, and it is also an effective source of information for people who generally verify information.

In the future, we intend to conduct experiments with a larger number of people using the system to investigate trends by age and changes in the system usage over time.

References

[1] Lerman, K. and Ghosh, R.: Information Contagion: an Empirical Study of the Spread of News on Digg and Twitter Social Networks, *Proc. 4th International Conference on Weblogs and Social Media* (*ICWSM-10*), pp.90–97 (2010).

[2] Let's discuss panic-buying at the supermarket the japan times (2020), available from ⟨https://www.japantimes.co.jp/life/2020/04/07/language/lets-discuss-panic-buying-supermarket/⟩ (accessed 2020-12-04).

[3] Coronavirus: Outcry after Trump suggests injecting disinfectant as treatment, BBC NEWS (2020), available from ⟨https://www.bbc.com/news/world-us-canada-52407177⟩ (accessed 2020-12-04).

[4] An Done on our continuity strategy during COVID-19 (2020), available from ⟨https://blog.twitter.com/en_us/topics/company/2020/An-Done-on-our-continuity-strategy-during-COVID-19.html⟩ (accessed 2020-12-04).

[5] Facebook, YouTube, and Twitter removed a COVID-19 video tweeted by Trump due to false coronavirus information—but millions had already viewed it, BUSINESS INSIDER (2020), available from ⟨https://www.businessinsider.com/facebook_youtube_twitter_take_down_false_covid_video_shared_trump_2020_7⟩ (accessed 2020-12-04).

[6] Panic-buying of 'made in China' tissues and toilet paper erupts in Japanese cities, the japan times (2020), available from ⟨https://www.japantimes.co.jp/news/2020/02/29/national/toilet-paper-tissue-coronavirus/⟩ (accessed 2020-12-04).

[7] Abstract of Information and Communication Media Usage Time and Information Behavior Survey Report 2019, Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications (2020), available from ⟨https://www.soumu.go.jp/iicp/research/results/media_usage-time.html⟩ (accessed 2020-12-04) (Japanese).

[8] Kawakami, Y.: *Social psychology of information propagation*, pp.32–47, SAIENSU-SHA (1977).

[9] Allport, G.W. and Postman, L.: *The psychology of rumor*, pp.41–60, Iwanami Shoten (2008).

[10] Knapp, R.H.: A Psychology of Rumor, *The Public Opinion Quarterly*, Vol.8, No.1, pp.22–37 (1944).

[11] Wang, L. and Fussell, S.R.: More Than a Click: Exploring College Students' Decision-Making Processes in Online News Sharing, *Proc. ACM on Human-Computer Interaction*, Vol.4, No.9, pp.1–20 (2020).

[12] Kakimoto, D., Miyabe, M. and Yoshino, T.: Proposal of Information credibility Verification behavior Facilitation System for Preventing False Rumors Spreading, *FIT2016*, D-029, pp.107–108 (2016) (Japanese).

[13] Gupta, A. and Kumaraguru, P.: Credibility Ranking of Tweets during High Impact Events, *Proc. 1st Workshop on Privacy and Security in Online Social Media*, No.2, pp.1–8, ACM (2012).

[14] Qazvinian, V., Rosengren, E., Radev, D.R. and Mei, Q.: Rumor has it: identifying misinformation in microblogs, *Proc. Conference on Empirical Methods in Natural Language Processing*, pp.1589–1599, ACM (2011).

[15] Lim, W.Y., Lee, M.L. and Wynne, H.: iFACT: An Interactive Framework to Assess Claims from Tweets, *Proc. 2017 ACM on Conference on Information and Knowledge Management*, pp.787–796, ACM (2017).

[16] Miyabe, M., Nadamoto, A. and Aramaki, E.: Construction of Services for Preventing False Rumors Spreading focusing on correcting information human mentioned, *Journal of Information Processing*, Vol.55, No.1, pp.563–573 (2014) (Japanese).

[17] Ennals, R., Trushkowsky, B. and Agosta, J.M.: Highlighting disputed claims on the web, *Proc. 19th International Conference on World Wide Web*, pp.341–350 (2010).

[18] Saito, F., Shoji, Y. and Yamamoto, Y.: Highlighting Weasel Sentences for Promoting Critical Information Seeking on the Web, *International Conference on Web Information Systems Engineering*, pp.424–440 (2019).

[19] Schwarz, J. and Morris, M.: Augmenting web pages and search results to support credibility assessment, *Proc. SIGCHI Conference on Human Factors in Computing Systems* (*CHI'11*), pp.1245–1254, ACM (2011).

[20] Gao, M., Xiao, Z., Karahalios, K.G. and Fu, W.T.: To Label or Not to Label: The Effect of Stance and Credibility Labels on Readers' Selection and Perception of News Articles, *Proce. ACM on Human-Computer Interaction*, Vol.2, No.55, pp.1–16 (2018).

[21] Information and Communications in Japan, Chapter4 Promotion of Inclusion through ICT, Section4 (2018), Information & Communications Statistics Database, Ministry of Internal Affairs and Communications, available from ⟨https://www.soumu.go.jp/johotsusintokei/whitepaper/h30.html⟩ (accessed 2020-12-04).

Ryota Nishimura was born in 1997. He received his B.E. degree from Wakayama University, Wakayama, Japan, in 2021. He is a Master's Student of Graduate School of Systems Engineering, Wakayama University. He is a member of IPSJ. His current interest is social applications of big data, especially preventing the spread of rumors.

Mai Miyabe Hirabayashi is a project assistant professor in Graduate School of Medicine, The University of Tokyo. She received her M.E. and Ph.D. degrees in Engineering from Wakayama University in 2008 and 2011, respectively. Her research topics include communication support and social media analysis.

Takashi Yoshino is a professor in Faculty of Systems Engineering, Wakayama University. He received his B.E. and M.E. degrees in electrical engineering from Kagoshima University in 1992 and 1994, respectively. He received Dr. of Information and Science from Tohoku University in 2001. His research interests include communication support and social applications of big data. He is a member of ACM, IEEE, IEICE.