# Calculating Cooking Recipe's Difficulty based on Cooking Activities

Kazuma Kusu
Graduate School of Culture and Information Science, Doshisha University
1-3 Tatara-Miyakodani
Kyotanabe, Kyoto 610-0394, Japan
kusu@ilab.doshisha.ac.jp

Nozomi Makino *
Faculty of Culture and Information Science, Doshisha University
1-3 Tatara-Miyakodani
Kyotanabe, Kyoto 610-0394, Japan
bin0092@mail4.doshisha.ac.jp

Takamitsu Shioi
School of Computing, Tokyo Institute of Technology
2-12-1 Ookayama
Meguro-ku, Tokyo 152-8550, Japan
shioi@de.cs.titech.ac.jp

Kenji Hatano
Faculty of Culture and Information Science, Doshisha University
1-3 Tatara-Miyakodani
Kyotanabe, Kyoto 610-0394, Japan
khatano@mail.doshisha.ac.jp

## ABSTRACT

In recent years, there have been plenty of cooking recipes on the internet. Because it has caused difficulty in searching for a cooking recipe that suites one's needs, many studies have researched on how to make these recipes easily accessible. Existing studies have focused on the components of cooking recipes, such as ingredients, nutrients, and condiments. However, if the recipe contains difficult cooking activities for the user, the user may not be able to cook the dish according to the recipe. Consequently, it is necessary to match the searched recipe's cooking skills. Therefore, in this work, we define four-levels of difficulty of cooking activities, and we propose a method for calculating a cooking recipe's difficulty level when searching for one that matches a user's cooking skills. Then, our approach was evaluated using a questionnaire survey on a crowdsourcing site. The results of our experiment showed that our approach can reduce the burden on users when searching for appropriate cooking recipes.

## CCS CONCEPTS

• **Information systems** → *Multimedia information systems*; *Users and interactive retrieval*; *Retrieval tasks and goals*; *Specialized information retrieval*; • **Human-centered computing** → *Accessibility design and evaluation methods*; • **Social and professional topics** → *Cultural characteristics*;

## KEYWORDS

Recipe Difficulty, Cooking Actions, Recipe Searching

*Currently, she belongs to KYOCERA Communication Systems Co., Ltd., Japan.


*CEA2017, Melbourne, Australia*

DOI: 10.1145/3106668.3106673

## 1 INTRODUCTION

The Internet has been growing at an unprecedented rate. Currently, people communicate via a specific site on the Internet. Using a specific site, people can access a topic of their interest and communicate with each other easily. This type of interaction is at times called an "online community." An online community can act as an information system where members can post, comment on discussions, offer advice, or collaborate. There are many types of online community sites that have emerged in recent years. Cookpad[1] had about 2.58 million cooking recipes and 1.92 million paying members as of March 2017 [2]. The increasing numbers of cooking recipes and paying members every year indicate a strong demand for an efficient way of searching for recipes online.

Conventional studies pertaining to cooking recipe search have focused on various components of a recipe, such as ingredients, nutrients, and condiments, to achieve an effective cooking recipe search. [4, 10]. A user's cooking skills are the most important factor in completing a cooking process, because a cooking recipe recommended by an online food community site may not be cooked properly if a user lacks cooking skills. In short, the most important aspect in searching for cooking recipes is to match a cooking recipe's difficulty level with a user's cooking skills. Therefore, this study aims to implement a recipe search system for identifying cooking recipes that match a user's cooking skills. Because the difficulty levels of cooking recipes have never been determined, we propose a method for calculating the difficulty level of each cooking recipe.

First, we define four levels of difficulty of cooking activities. Second, we propose a method for calculating a cooking recipe's difficulty level to match a user's cooking skills. To evaluate our approach, we conduct a questionnaire survey on a crowdsourcing site.

[1]Cookpad's home page in Japan. https://cookpad.com/(accessed May 8, 2017)
[2]Cookpad Presentation Material for 4Q of FY 2016 https://cf.cpcdn.com/info/assets/wp-content/uploads/20170228100045/2016Q4en.pdf (accessed May 8, 2017)

## 2 RELATED WORK

Several researchers have studied scoring methods for recipe searching that satisfy a user's information needs. These studies are within the realm of the field of information retrieval. The most typical approach used by these studies concerns various components of a recipe, such as ingredients, nutrients, and condiments, in calculating a score for each recipe. Akazawa et al. proposed a recipe search system that considers ingredients in a refrigerator [1]. His study examined the remaining amount, expiration date, and last used date of ingredients in the refrigerator for calculating recipe scores and then arranged the scores in descending order. This approach enables users to search for recipes effectively because users can find ingredients in the refrigerator that must be used as soon as possible.

Ueda et al. proposed a recommendation method based on a user's cooking history [11]. In their study, they extracted the personal preferences of a user from the recipes that he/she had cooked, and the identified the specificity of the ingredients. The characteristics of the specificity were expressed as FF-IRF (Foodstuff Frequency Inverted Recipe Frequency), which was derived from a term weighting scheme named TF-IDF (Term Frequency Inverted Document Frequency) in the research field of information retrieval. Using FF-IRF, they calculated the score of a recipe based on a user's taste, and arranged the scores in descending order. In short, recipes using a frequently used foodstuff that is not used in other recipes are recommended.

There is another type of recipe scoring method based on cooking activities and the recipe's ingredients and condiments. Yajima et al. implemented a recipe recommendation system using the cooking-procedure-based scoring method, and suggested cookable recipes for certain users [12]. In this scoring method, they extracted verbs from cooking procedures in recipes, and defined discrete values with respect to each verb in the recipes. The discrete values, which are called "cooking method score" in [12], denote the difficulty level of a cooking activity; however, this definition is based on the researchers' subjective view.

Iwamoto et al. also proposed a scoring method that considers not only the difficulty level of a cooking activity but also the length of a recipe [3]. Yajima et al. defined the difficulty level of cooking activities based on the researchers' subjective view, whereas Iwamoto et al. defined the difficulty level using an examination of cooking activities in home economics, which is certificated by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The difficulty level of each cooking activity was defined by the criteria of a four-step-level examination; certain cooking activities not defined in the criteria were defined by the researchers[3].

## 3 PROPOSED METHOD

As described in Section 2, it is difficult to define the difficulty level of recipes in an objective way using the criteria of the certificate examination. In the criteria, only cooking activities covered by the examination are defined. Consequently, several verbs pertaining to cooking activities were not defined, and their scores were not also calculated. In this study, we, first, define the difficulty level of cooking activities extracted from all recipes in Japanese schoolbooks of home economics, which are certificated by MEXT. The textbooks of the certificate examination are not used in this study because 1) there is no recipe in the textbooks, and 2) the cooking activities in recipes are needed to define subjectively the difficulty level of recipes.

In this section, the schoolbooks of home economics, ways to extract text data from the schoolbooks, and ways to calculate and define the difficulty levels of cooking activities and recipes are explained.

### 3.1 Text Data in Schoolbooks

To calculate the difficulty level of cooking activities, a criterion for the difficulty level of recipes must be defined, as is the case with [3]. However, the criteria defined in [3] are inadequate in terms of qualitative data. Hence, we constructed the criteria using cooking activities extracted from all recipes in Japanese schoolbooks on home economics. The schoolbooks are certificated, in accordance with a guideline compiled by MEXT. Hence, their appropriateness is guaranteed [4]. In short, the cooking activities that appeared in the schoolbooks are objective, as these schoolbooks are certainly used in the 12-year curriculum of Japan's school education.

This study employed 24 schoolbooks in total (two schoolbooks for elementary school, three for junior high school, 17 for high school, and two for specialized vocational high school), based on a list of schoolbooks compiled by MEXT in fiscal 2015 [5,6,7]. As a result, 919 recipes were extracted from the schoolbooks (30 from schoolbooks for elementary schools, 131 from schoolbooks for junior high schools, 478 from schoolbooks for high schools, and 280 from schoolbooks specialized vocational high schools).

### 3.2 Extraction of Cooking Activities

After extracting recipes from the schoolbooks, we use a text processing tool called KyTea [8] to extract cooking activities from the cooking instructions in recipes. KyTea enables us to segment a text into words, annotated with parts-of-speech and pronunciations [6, 8]. Moreover, named entities of a recipe can also be recognized using PWNER [9] [7]. PWNER enables us to automatically assign the tags shown in the Table 1 to each named entity of a recipe [9].

After segmenting the words and recognizing the named entities of a recipe, words with the tag "Ac" were extracted, indicating the cooking activities required for the cooking instructions.

[3]Currently, this certificate is a six-step-level examination, after it was revised.

[4]The guideline is a regulation enabling students around Japan to be equally educated based on a certain standard stated by Japan's School Education Act.

[5]a list of schoolbooks for elementary schools: http://www.mext.go.jp/a_menu/shotou/kyoukasho/mokuroku/27/__icsFiles/afieldfile/2015/04/21/1357046_1_1.pdf (accessed May 8, 2017)

[6]a list of schoolbooks for junior high schools: http://www.mext.go.jp/a_menu/shotou/kyoukasho/mokuroku/27/__icsFiles/afieldfile/2015/07/15/1357046_2.pdf (accessed May 8, 2017)

[7]a list of schoolbooks for high schools: http://www.mext.go.jp/a_menu/shotou/kyoukasho/mokuroku/27/__icsFiles/afieldfile/2015/07/15/1357046_3.pdf (accessed May 8, 2017)

[8]KyTea: http://www.phontron.com/kytea/index-ja.html (accessed May 8, 2017)

[9]a manual of recipe text processing: http://www.ar.media.kyoto-u.ac.jp/how-to/recipe-NLP/ (accessed May 8, 2017)

**Table 1: Tags of named entities in a recipe**

| Tag | Meaning | Explanation |
|---|---|---|
| F | Ingredients | something to be able to eat, dump, change its quantity including an intermediate product |
| T | Tools | a tool is a physical entity |
| Ac | Cooking Activities | the stem of a word (its subject is cook) |
| Af | transformation of an ingredient | the stem of a word (its subject is ingredient) |
| Sf | State of Ingredient | representation of an ingredient's aspect before and after change |
| St | State of Tools | representation of a tool's aspect before and after change |
| D | Duration Time | the duration time for cooking |
| Q | Quantity | the quantity of an ingredient |

## 3.3 Difficulty Level of Cooking Activities

As we described in Section 3.1, we utilized the schoolbooks used in four different levels of education. Accordingly, the difficulty of cooking activities has four levels. Therefore, if a cooking activity Ac is given, the difficulty $D_a$ can be calculated using the following equation:

$$\{D_a(a) \mid a \in A_r, r \in R\} = \begin{cases} 1 & (\mathtt{if}\ a \in E) \\ 2 & (\mathtt{if}\ a \in J - E) \\ 3 & (\mathtt{if}\ a \in H - J - E) \\ 4 & (\mathtt{if}\ a \in S - H - J - E) \end{cases} \tag{1}$$

where $R$ is a set of recipes; $A_r$ is a set of cooking activities in a recipe $r \in R$; E is a set of cooking activities extracted from the schoolbooks of elementary school, J from those of junior high school, H from those of high school, and S from those of specialized vocational high school. The difficulty of cooking activities is defined as $D_a(a)$ because the more we learn $a$ during our primary education, the lower $D_a(a)$ is.

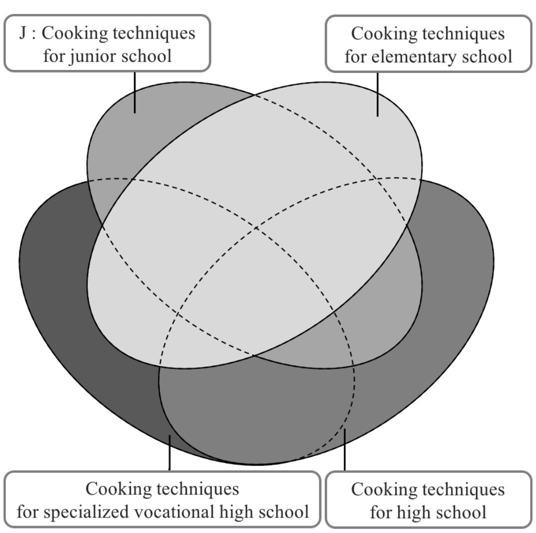


**Figure 1: The Euler diagram represents our definition of the difficulty levels of cooking activities.**

As a result, we could extract 485 cooking activities: level 1 contains 92 cooking activities; level 2, 152; level 3, 135; and level 4, 106, as shown in Figure 2. The examples of cooking activities in each level are shown below.

**difficulty-level 1:** 温める (heat), 洗う (wash), いちょう切り (quarter-rounds), 輪切り (rounds)
**difficulty-level 2:** 泡立てる (whisk), くし形切り (wedges), 裏返す (turn over)
**difficulty-level 3:** 揚げる (fly), 裏ごし (strain), 縛る (tie)
**difficulty-level 4:** あられ切り (dice), 面取り (chamfer), フランベ (flambé)

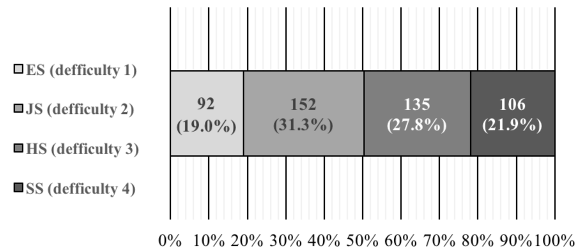


**Figure 2: A rate of difficulty-levels**

## 3.4 Difficulty Level of Recipes

In this section, we propose eight kinds of scoring methods for calculating the difficulty level of cooking activities using our definition. In addition to our proposed method, we used a scoring method proposed by Ueda et al. [11]. We also describe our scoring methods for the difficulty level of a recipe. For ease of comprehension, let us suppose that given a recipe with a cooking procedure containing ten cooking activities, including three that are defined as difficulty-level 1, two that are difficulty-level 2, three that are difficulty-level 3, and two that difficulty-level 4.

**Method 1, Method 2:**
Method 1 and Method 2 use a mode indicating the number of difficulty-levels for determining the difficulty level of a recipe. The difficulty of Method 1 $D_r^1$ is calculated by Formula (2), and one of Method 2 is calculated by Formula 3.

$$D_r^1 = \min(d_{mode}) \tag{2}$$

$$D_r^2 = \max(d_{mode}) \tag{3}$$

We propose two cases based on a calculation of a minimum and a maximum, because the mode of the number of cooking activities in each difficulty-level $d_{mode}$ is not a unique solution. For the above example, there are two modes, namely, difficulty-levels 1 and 3 in a cooking procedure. In this case, Method 1 selects difficulty-level 1, and Method 2, difficulty-level 3.

**Method 3, Method 4:**
Methods 3 and Method 4 use a median indicating the number of difficulty-levels for determining the difficulty of a recipe.

$$D_r^3 = \lfloor \mathrm{median}(d_{B_r}) \rfloor \tag{4}$$

$$D_r^4 = \lceil \text{median}(d_{B_r}) \rceil \quad (5)$$

where $B_r$ is a set of cooking activities in a recipe $r$, and $d_{B_r}$ is a set of the difficulty-levels of every $B_r$. For the above example, a total of cooking activities is ten, which is an even number.Hence, a median has possibilities to be a real number. In this case, a median is rounded down in Method 3, and rounded up in Method 4.

**Method 5:**

This method uses a maximum of difficulty-level as a difficulty of a recipe, based on the following formula:

$$D_r^5 = d_{max} \quad (6)$$

where $d_{max}$ is the difficulty-level of maximum number of cooking activities in a recipe $r$. For the above example, the difficulty level of a recipe is difficulty-level 4 because this recipe contains all the difficulty-levels in its cooking procedure.

**Method 6:**

This method uses a mean of all difficulties of cooking activities in a cooking procedure of a recipe, as follows:

$$D_r^6 = \frac{1}{n_r} \sum_{d=1}^{4} d \cdot n_d \quad (7)$$

where $n$ is the number of the cooking activities in a recipe $r$, and $n_d$ is the number of cooking activities in each difficulty-level $d$. For the above example, the difficulty level of a recipe equals 2.4, which is calculated as follows:

$$(1 * 3 + 2 * 2 + 3 * 3 + 4 * 2)/10 = 24/10 = 2.4 \quad (8)$$

**Method 7:**

This method uses each difficulty-level as a bias. For example, let us suppose that $d$ is a difficulty level in cooking activities. The difficulty level of a recipe $S$ is calculated as follows Formula (9) in Method 5.

$$D_r^7 = \sum_{d=1}^{4} \left( \frac{n_d}{N} \cdot d \right) \quad (9)$$

In Formula (9), $N$ means the total of cooking activities, and $n_l$ means the number of difficulty-level $l (\in L)$ in a recipe.

**Method 8:**

Method 8 uses the frequency of each cooking activity in all recipes. Given a set of the cooking activity as $B$, the number of cooking activity $b (\in B)$ in all recipes as $N_{ct}^b$, the number of cooking activity $b$ in a recipe as $N_{ct}^b$, the number of cooking activities in a recipe as $n_{ct}^B$, and the number of similar difficulty-level with $b$ in all recipes as $N_{ct}^{l_b}$, this method's score is calculated by using the following Formula (10).

$$D_r^8 = \sum_{b \in B} \frac{\left( \left( 1 - \frac{N_{ct}^b}{N_{ct}^{l_b}} \right) \cdot l_b \cdot n_{ct}^b \right)}{N_B} \quad (10)$$

**Method 9:**

Method 9 is proposed by Ueda et al. [11], which enables the calculation of the specificity of each cooking activity. Given $N_r$ means the number of recipes and $N_r^b$ means the number of recipe including cooking activity $b$, $iRf^b$ which is defined as $-log_{10}(N_r^b/N_r)$. Therefore, a score of a recipe is calculated by the following Formula (11).

$$D_r^9 = \sum_{b \in B} \left( \frac{n_{ct}^b}{n_B} \cdot IRF^b \cdot l_b \right) \quad (11)$$

## 3.5 Measurement of a Recipe's Difficulty Level

We described nine scoring methods for calculating the difficulty level of a recipe in Section 3.4. In this section, we introduce thresholds to determine the four levels of difficulty for scoring methods that produced a real-valued score.

We used directly a score produced by Method 1, 2, 3, 4, and 5 as the difficulty level of a recipe, because these produced four levels of difficulty.

Methods 6, 7, 8, 9 produced a real-valued score from 0 to 4. Hence, when we determine the four levels of difficulty from a non-normal distributed real-valued score, we set three thresholds by using a median (m) and the quartile deviation (q). The difficulty level of a recipe is difficulty-level 1 if a score is in $[min, m - q)$, 2 if a score is in $[m - q, m)$, 3 if a score is in $[m, m + q)$, and 4 if a score is in $[m + q, max]$.

# 4 EXPERIMENTAL EVALUATION

In this section, we explain an experiment for evaluating each scoring method proposed in Section 3.4. Here, we calculated an accuracy of the number of correct judgment using our method to that of given recipes. Moreover, for evaluating the validness of our defined difficulty levels of cooking activities, we used real recipe data on a food online community site, calculated their difficulties, and the concordance rate for judgment of our approach and a third person's view.

## 4.1 Evaluating the Difficulty Level of Recipes

Using the following equation, we calculated the accuracy of the number of correct judgment using each method to that of given recipes:

$$\frac{1}{4} \sum_{d \in D} \frac{n_d}{N_d} \quad (12)$$

where $D$ is a set of the difficulty-level of a recipe; $N_d$ is the number of recipes extracted from schoolbooks of elementary school if $d$ equals 1, junior high school if $d$ equals 2, high school if $d$ equals 3, and specialized vocational high school if $d$ equals 4. In addition, $n_d$ is the number of recipes with correctly determined difficulty level.

**Table 2: Accuracies of Our Difficulty of Recipes**

| | Method 1 | Method 2 | Method 3 |
|---|---|---|---|
| accuracy | 0.044 | 0.044 | 0.041 |
| | **Method 4** | **Method 5** | **Method 6** |
| accuracy | 0.042 | 0.572 | 0.370 |
| | **Method 7** | **Method 8** | **Method 9** |
| accuracy | 0.368 | 0.362 | 0.332 |

Table 2 shows that accuracy of Method 5 seems to be the highest among all scoring methods. Therefore, conducting a hypothesis testing for the difference in the population proportions between Method 5 and others, its result indicates statistically significant difference under the significant level of 1%. As a result, we believe that using Method 5 is best for calculating the difficulty level of recipes.

## 4.2 The Cover Rate for the Cooking Activities

In this section, we describe another experiment for confirming the usefulness of our method for determining the difficulty level of cooking activities. In this experiment, we used the recipes on a food online community site named Cookpad dataset[10], because this site consists several recipes that are publicly available in Japan.

From this dataset, we take some recipes for keeping the randomness. In short, if we set the confidence level $\alpha$ to 5%, the minimum sample size is 384 recipes calculated by the following formula [2]:

$$n = Z_{\frac{\alpha}{2}}^2 \cdot \frac{p \cdot (1-p)}{\alpha^2} = 1.96^2 \cdot \frac{0.5 \cdot (1-0.5)}{0.05^2} \approx 384 \qquad (13)$$

where $Z_{\frac{\alpha}{2}}$ is the $\alpha$% of $z$-score on the standard normal distribution, and $p$ is the ratio of answer. Here, we set the value of $p$ to 0.5, because Formula (7) produces a maximum value of $n$.

As stated above, we extracted 400 recipes in round figures as a sample size by conducting a random sampling from 1.72 million recipes in Cookpad's dataset space.Moreover, we extracted cooking activities from sample recipes and calculated a percentage $C$ of cooking activities to determine the difficulty levels (cover rate) using the following formula:

$$C = \frac{n_a}{N_a} \qquad (14)$$

In Formula (14), $N_a$ is the number of cooking activities in the sample recipes belonging to a-th difficulty level, and $n_a$ is the number of cooking activities that determined the difficulty-level of sample recipes. After ten random samplings, the average of $C$ was 84% (max: 85%, min: 83%).

## 4.3 Questionnaire in the Experiment

To evaluate the validness of our method in determining the difficulty level of cooking activities, we also conducted a questionnaire survey.Figure 3 shows a sample image of our questionnaire.

We aimed to confirm whether respondents could recognize the difference of each difficulty-level of recipes. Hence, we designed a questionnaire comprising four different levels of difficulty of recipes, and two questions on whether four recipes are listed in the order of their difficulty or randomly. However, it is difficult for respondents to determine appropriately the difficulty of some recipes owing to a variety of kinds of recipes available in real life. In this experiment, therefore, we only used recipes under a given category of dish in Cookpad's dataset to determine easily the difficulty level of the recipes.

Using Method 5 described in Section 3.4, we obtained 729 categories containing all difficulty-levels of at least one recipe. Finally, we selected 100 categories whose category names were cooking names listed in the questionnaire survey.

[10]Cookpad dataset: http://www.nii.ac.jp/dsc/idr/cookpad/cookpad.html (accessed May 8, 2017)

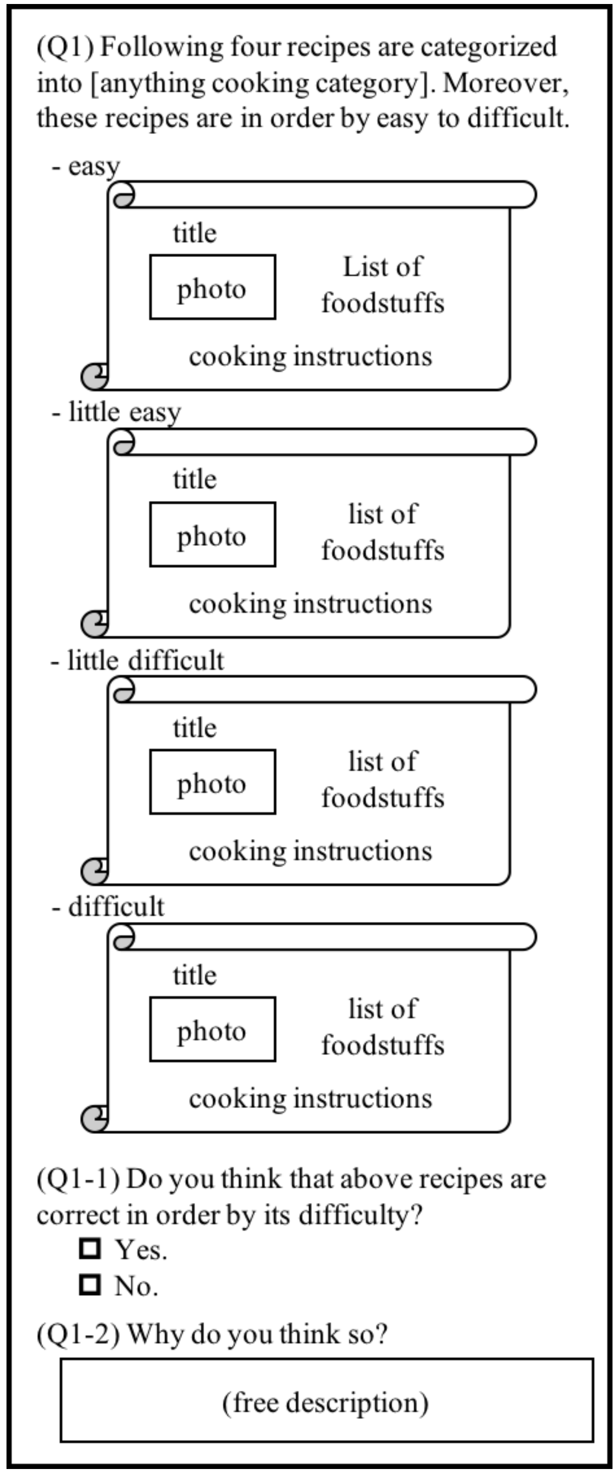


**Figure 3: Construction of a Questionnaire**

In creating the questionnaire, we extracted two sets of four different difficulty-levels of recipes from a certain category, and created two cases of recipes in the order of their difficulty levels and randomly. We created 200 questionnaires, as described above, and conducted questionnaire investigation by using the crowdsourcing site Lancers[11] in Japan. In this questionnaire, we decided that the number of respondents is should be 140 (40 males and 100 females), because Cookpad reported that its paying members are largely comprised of women.

### 4.4 Experimental Results and Consideration

The result of the questionnaire survey is shown in Table 3. In Table 3, we investigated the degree of coincidence between the actual and the respondents' answers.

**Table 3: A Result of Questionnaires**

| | | actual answer | | |
|---|---|---|---|---|
| | | ordered | not ordered | total |
| respondent's answer | ordered | 89 (44.5%) | 3 (1.5%) | 92 |
| | not ordered | 11 (5.5%) | 97 (48.5%) | 108 |
| total | | 100 | 100 | 200 |

Here, we conducted a hypothesis testing for the difference in the population proportions between the rate of ordered and not ordered answers. This null hypothesis ($H_0$) is $p = p_0$, and this alternative hypothesis ($H_1$) is $p \neq p_0$. Moreover, we calculated the test statistic $z$ using the following equation:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \tag{15}$$

where $\hat{p}$ is a sample proportion, $p_0$ is the hypothesize population proportion, and $n$ is the sample size. Then, the $p$-value is lower than 0.01 in this test, so that the ratio difference between the "ordered" and "not ordered" answer gives a statistically significant difference. The result shows that our approach could determine the difficulty level of recipes that match a user's needs.

Meanwhile, we listed the respondents' reasons for feeling that a recipe is difficult.

- a detailed cooking instructions
- no photograph and only texts
- too many foodstuffs
- too long cooking time

Therefore, the respondents reasons show that the level of difficulty of cooking is determined by not only the cooking activities but also the duration of cooking and the number of words, photographs, and amount of foodstuffs listed in a recipe.

## 5 CONCLUSION

In this work, we aimed to implement a recipe search system that could effectively identify recipes that match a user's cooking skills.We defined the difficulty level of the cooking activities extracted from Japanese schoolbooks in home economics, which are compiled by MEXT. Moreover, we proposed a scoring method for calculating the difficulty level of a recipe based on the cooking activities extracted from these schoolbooks.

In the experimental evaluation, we conducted a questionnaire survey using a crowdsourcing site to confirm that our approach could correctly determine the difficulty level of a recipe and match the difficulty level of a recipe with a user's cooking skills. As a result, our approach achieved an accuracy of 89.5% on average.

Future studies must focus on not only the cooking activities but also the duration of cooking and the number of words, photographs, and foodstuffs listed in a recipe. Moreover, our definition of the level of difficulty of cooking activities cannot determine that of unknown words. Hence, it is necessary to use a cooking ontology or examine ways to extract synonyms to improve the cover rate. Finally, implementing a recipe search system that enables an efficient search of recipes that suit a user's skills is important.

## ACKNOWLEDGEMENT

We would like to thank Tetsuro Sasada of Kyoto University for his advice on how to learn a model of KyTea. This work was supported by JSPS KAKENHI Grant Number JP15H02701. Moreover, we used recipe data provided by Cookpad and the National Institute of Informatics.

## REFERENCES

[1] Yasuyuki Akazawa and Hisashi Miyamori. 2012. Reizouko Shokuzai wo kouryo shita Ryouri Resipi Kensaku Sisutemu no Teian (Proposal of a Search System for Cooking Recipes Considering Ingredients in Refrigerators). In *Proceedings of DEIM Forum 2012.*

[2] Pascal Ardilly and Yves Tillé. 2006. *Sampling Methods: Exercises and Solutions.* Springer New York.

[3] Junya Iwamoto and Hisashi Miyamori. 2012. Tyori no Nan-ido wo Kouryo shita Resipi Kensaku Sisutemu no Teian (Proposal of a Search System for Cooking Recipes Considering Difficulty of Cooking). In *Proceedings of DEIM Forum 2012.*

[4] Shihono Karikome and Atsushi Fujii. 2010. A System for Supporting Dietary Habits: Planning Menus and Visualizing Nutritional Intake Balance. In *Proceedings of the 4th International Conference on Uniquitous Information Management and Communication (ICUIMC '10).* Association for Computing Machinery, 56:1–56:6.

[5] Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, and Tetsuro Sasada. 2014. Flow Graph Corpus from Recipe Texts. In *Proceedings of the Nineth International Conference on Language Resources and Evaluation.* 2370–2377.

[6] Shinsuke Mori and Graham Neubig. 2011. A Pointwise Approach to Pronunciation Estimation for a TTS Front-end. In *Proceedings of the InterSpeech2011.* Florence, Italy, 2181–2184.

[7] Shinsuke Mori, Tetsuro Sasada, Yoko Yamakata, and Koichiro Yoshino. 2012. A Machine Learning Approach to Recipe Text Processing. In *Proceedings of the 1st Cooking with Computer Workshop.* 29–34.

[8] Graham Neubig, Yosuke Nakata, and Shinsuke Mori. 2011. Pointwise Prediction for Robust, Adaptable Japanese Morphological Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics.* 529–533.

[9] Tetsuro Sasada, Shinsuke Mori, Tatsuya Kawahara, and Yoko Yamakata. 2016. *Named Entity Recognizer Trainable from Partially Annotated Data.* Springer Singapore, 148–160.

[10] Hidekazu Shiozawa and Yusuke Mitamura. 2007. Shokuzai no Yuusendo wo kouryo shita Ryouri Resipi no Kensaku (Using Priorities of Ingredients for Interactive Cooking Recipe Search). *IPSJ SIG technical reports* 2007, 41 (2007), 51–57.

[11] Mayumi Ueda, Mari Takahata, and Shinsuke Nakajima. 2011. User's Food Preference Extraction for Personalized Cooking Recipe Recommendation. In *Proceedings of the Second International Conference on Semantic Personalized Information Management: Retrieval and Recommendation - Volume 781 (SPIM'11).* CEUR-WS.org, 98–105.

[12] Asami Yajima and Ichiro Kobayashi. 2009. "Easy" Cooking Recipe Recommendation Considering User's Conditions. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 03 (WI-IAT '09).* IEEE Computer Society, 13–16.

[11] Lancers' website, http://www.lancers.jp/ (accessed May 8, 2017)