Recipe Recommendation Method by Considering the User's Preference and Ingredient Quantity of Target Recipe

Mayumi Ueda, Syungo Asanuma, Yusuke Miyawaki, and Shinsuke Nakajima

***Abstract*—There are many websites and researches that invoke cooking recipe recommendation. However, these websites present cooking recipes on the basis of entry date, access frequency, or the recipe's user ratings. They do not reflect the user's personal preferences. We have proposed a recipe recommendation method that is based on the user's food preferences. For extracting the user's food preferences, we use his/her recipe browsing and cooking history. In our previous work, we consider only existence of non-existence of each ingredient in the cooking recipe for extracting the preferences. In order to reflect the truly user's preferences, we propose a scoring method of cooking recipes based on user's food preferences and the quantity of the ingredient in a recipe.**

***Index Terms*—cooking recipe recommendation, user's food preferences, preference extraction, cooking and browsing history, quantity of ingredient in recipe.**

I. Introduction

As a result of the lifestyle-related disease epidemic, dietary life is now attracting attention. Good eating habits are important for maintaining a healthy life. However, menu planning required one to take various factors into consideration, such as the nutritional values, food in stock, food preferences, and cost. Thus, people need to expand a lot of effort toward planning their daily menu. Against this background, a number of cooking websites comprising various food recipes have been launched, such as Cookpad[1] and Rakuten Recipe[2]. Many people refer to these websites when planning their menu. Cookpad contains one and a half million recipes and 20 million monthly users[3]. This data reflects the high demand for recipe-providing services. However, these websites do not reflect user's preferences and conditions, although these two factors need to be considered if the goal is to provide high satisfactory recipes.

Furthermore, several researches on cooking recipe recommendation for menu planning support have been conducted in the past. Mino et al. propose a method that takes the user's schedule into consideration[4]. This paper defines the evaluation value of either the intake or consumption calories that are assigned to each event in the user's schedules. Karikome et al. propose a system that helps users plan nutritionally balanced menus and visualize their dietary habits[5]. Their system calculates the nutritional value of each dish, and records this information in the form of a dietary log. Next, the system recommends recipes foster sound nutrition. Freyne et al. show the results of their investigation in which they compare three recommendation strategies: content-based, collaborative, and hybrid[6].

In these circumstances, we have proposed a recipe recommendation method based on the user's food preferences[7]. Our method breaks recipes down into their ingredients, and scores them on the basis of the frequency of use and specificity of the ingredients. Furthermore, our proposed system does not recommend dishes that are similar to the food the users have eaten over the past few days on the grounds that people do not want to eat similar dishes iteratively. Moreover, our system does not require any particular action on the user's past reflect his/her food preferences: it estimates the user's food preferences automatically through his/her recipe browsing and cooking history. Furthermore, we consider that our previous work could not reflect user's preferences completely, so we propose a new recipe recommendation method based on user's food preferences and the quantity of the ingredients. In this paper, we present a method for extracting the user's preferences, and reflecting the preferences based on the quantity of each ingredients.

This paper is structured as follows. Section II describes the method of scoring recipes and extracting user's preferences. Section III shows experimental results. Section IV shows concludes the paper.

II. Scoring Recipes and Extracting User's Preferences

In the recent years, concern over various health issues, such as lifestyle-related diseases and diets, has been growing. It has also been noted that picky eating is one of the main reasons causing these health issues. However, people do not want to eat food that they dislike even if it perfectly addresses their nutritional needs. They hope to derive essential nutrition solely from their favorite foods. Therefore, we try to extract user's food preferences.

A. Preferences for Ingredients

We express the user's food preferences I_k by using in the form of the following Eq(1).

$$I_k = I_k^+ - I_k^- \tag{1}$$

1) User's Favorite Ingredients: Fig.1 shows the key idea behind estimating user's favorite ingredients by his/her cooking history. Our method considers the ingredient that the

Manuscript received January 8, 2014; revised January 30, 2014. This work was supported in part by the MEXT Grant-in-Aid for Scientific Research(C) (#23500140).

M. Ueda is with University of Marketing and Distribution Sciences, Kobe, Japan. e-mail: Mayumi_Ueda@red.umds.ac.jp

S. Asanuma, Y. Miyawaki, and S. Nakajima are with Kyoto Sangyo University, Kyoto, Japan. email: {g1044002, i1258140, nakajima}@cc.kyoto-su.ac.jp

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

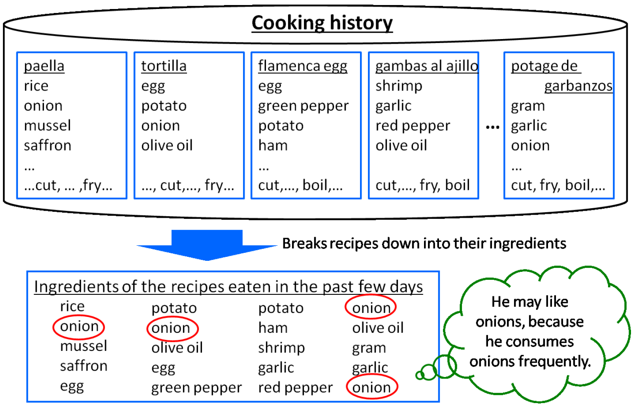

Fig. 1. Extracting the Favorite Ingredients Using Cooking History

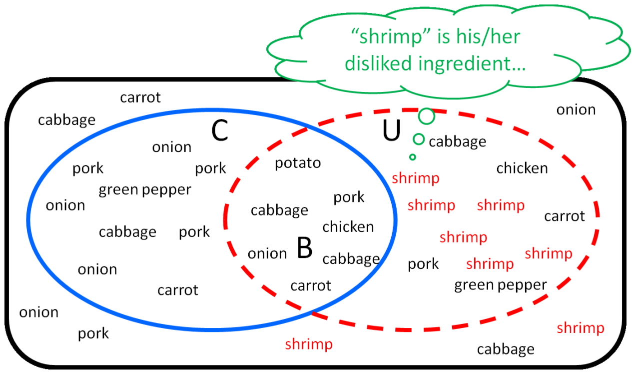

Fig. 2. Extracting the Disliked Ingredients Using Browsing and Cooking History

user eats repeatedly as his/her favorite ingredients. It breaks recipes down into their ingredient as the outset and calculated *the score of ingredients* I_k^+ by incorporating the frequency of use of the ingredients in the dishes that the target user has eaten (FF_k *: foodstuff frequency*) as well as the specificity of ingredients (IRF_k *: inverted recipe frequency*) into Eq(2). This equation is based on the idea of TF-IDF.

$$I_k^+ = FF_k \times IRF_k \qquad (2)$$

For estimating the user's favorite ingredients by using *the frequency of use of ingredient k* (FF_k), we utilize *the simple frequency of use of ingredient k* (F_k) during a definite period D, as shown in Eq(3).

$$FF_k = \frac{F_k}{D} \qquad (3)$$

Then, we calculate *the specificity of ingredient k* (IRF_k) using *the total number of recipe* (M) and *the number of recipes that contain ingredient k* (M_k), as shown in Eq(4).

$$IRF_k = \log \frac{M}{M_k} \qquad (4)$$

2) User's Disliked Ingredients: We consider that the user's food preferences are also influenced by his/her disliked ingredients. We estimate the user's disliked ingredients, by considering the ingredients in the recipes that he/she has never cooked, even if he/she has browsed the recipe details. Fig.2 shows the estimating method for user's disliked ingredients through the user's recipe browsing and cooking history. N corresponds to the set of ingredients in the recipes that the user has not browse. C corresponds to the set of ingredients in the recipes that the user has cooked over the past few days. U corresponds to the set of ingredients in the recipes that the user has not cooked, even if he/she has browse them completely. For example, "shrimp" in Fig.2 corresponds to the user's disliked ingredient. We calculate *the score of disliked ingredient* $k(I_k^-)$ in Eq.(5)).

$$I_k^-(x) = \begin{cases} 0 \ (0 < \frac{2|U_k|}{|A_k|} \leq 0.5) \\ (\frac{2|U_k|}{|A_k|} - 1)^x \ (0.5 < \frac{2|U_k|}{|A_k|} \leq 1) \end{cases} \qquad (5)$$

$|U_k|$ denotes the presence of ingredient k in U and A_k denotes the presence of ingredient k in the recipe that the user has browse them completely. We investigate the ratio or frequency of the user's avoidance of the ingredient that he/she dislikes, because he/she will use the ingredients that he/she does not like. x denotes the ratio or frequency of avoiding the ingredients, and we plan to verify x through some preliminary experiments.

B. Quantity of Ingredient in Recipe

Fig.3 shows the basic concept of our recipe recommendation method. The left figure shows the our concept of our previous work based on the user's food preferences. The method estimates a user's preferences from his/her past actions, such as through their recipe browsing and cooking history. And the right figure shows the new method based on the user's food preferences and the quantity of the ingredients in the recipe. Consider $RecipeA$ and $RecipeB$ for example (Fig.3). Recipe A contains 300 grams of pork, 60 grams of cabbage and 100 grams of potato, and Recipe B contains 50 grams of pork, 400 grams of cabbage and 20 grams of potato.

Our previous work scores recipe based on the user's food preferences from his/her past actions, such as recipe browsing and cooking history. For extracting the food preferences, our previous method breaks recipes down into their ingredients. That is, the score of recipes are determined by the contain or not contain of the ingredients. Because of both recipe contains the same ingredients, both recipes are same scores, even if he/she dislikes potato (left figure of Fig.3). Both recipes are same scores, even if he/she likes pork. This idea does not suit the eating habits.

Therefore, we propose a new scoring method based on the user's food preference and the quantity of the ingredients in the recipe. This method scores recipe based on the user's food preferences and the quantity of the ingredients in the recipe. That is, the scores of recipe that contains large amounts of favorite ingredients are high, even if both recipes contains same ingredients. As shown in right figure of Fig.3, even if both recipes contain same ingredients, the scores are different. If he/she likes pork, the score of Recipe A is higher than Recipe B, even if both recipes, even if both recipes are made up of the same ingredients. And if he/she dislikes potato, the score of Recipe B is higher than Recipe A.

For this method, we should not score recipes based on simple method, such as according to the quantity of each ingredients. Consider "100 grams of ingredient X" for example. We feel 100 grams of "pepper" is large quantity. However, we

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

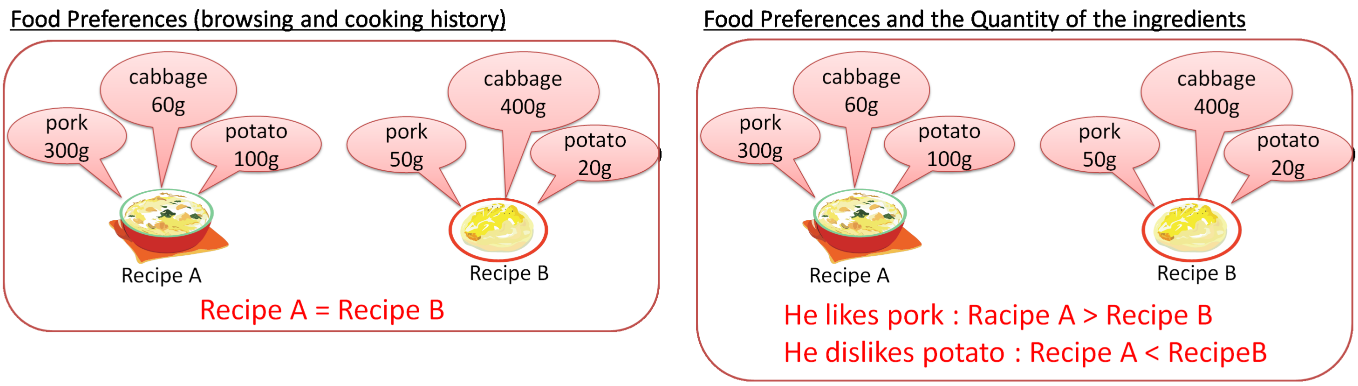

Fig. 3. The basic Concept of Our Previous Scoring Method (left) and new Scoring Method (right).

feel 100 grams of ”potato” is not so large. Hence, we propose the scoring method based on the standard quantity of each ingredients, and the dispersion quantity of each ingredients.

Fig.4 shows our basic idea of the scoring method for three ingredients, according to the average and dispersion quantity of each ingredients. The meanings of ”30 grams of ingredient” are different through three ingredients. As shown in middle figure of Fig.4, ”30 grams of ingredient p” is quite small quantity. Ingredient p have a relatively small effect on this recipe. On the other hand, ”30 grams of ingredient q” is quite large quantity for the ingredient. q Ingredient q have a significant impact for this recipe.

Therefore, we use the dispersion quantity of the ingredient for recipe scoring, that is calculated by the positioning of each ingredients in the all recipes.

C. Recipe Scoring

1) User's Food Preferences: Our method scores cooking recipes in accordance with the estimation results regarding favorite/disliked ingredients, and then provides recipes in decreasing order of the scores. In general, people do not like eating dishes similar to those they have eaten in the past few days. Therefore, our method weights recipes to avoid the repetition of similar dishes. *The score of cooking recipes* are defined as shown in Eq.(6).

$$Score(R) = \sum_{k \in R} I_k - \alpha \sum_{d=1} (w_d \cdot sim(R, R_d)) \quad (6)$$

d denotes the weight for avoiding repeating similar dishes iteratively. $sim(R, R_d)$ denotes the similarities between the considered recipe R and the recipe of the dish eaten d days ago R_d. The weight w_d for avoiding similar dishes eaten d days ago is defined as shown in Eq.(7).

$$w_d = 1 - \frac{d-1}{7} \quad (1 \leq d \leq 7) \quad (7)$$

2) The Quantity of the Ingredients: Our method calculates the weights of ingredient k using the average and standard deviation of the quantity of each ingredients. The average quantity of each ingredients are calculated by the average quantity of all recipes in the database. We calculate the standard deviation in Eq.(8).

$$\sigma_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (g_{k(i)} - \overline{g_k})^2} \quad (8)$$

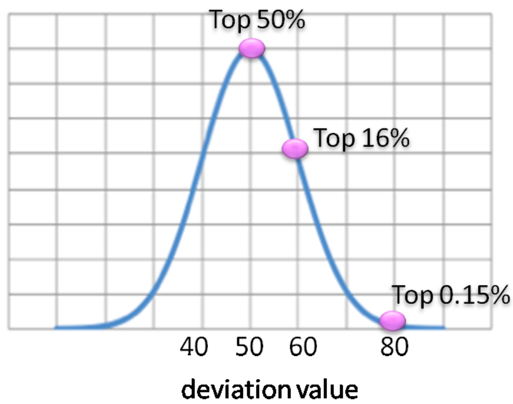

Fig. 5. The distribution of the deviation value for the weight

n denotes the number of recipes that contain ingredient k, $g_{k(i)}$ denotes the quantity of the ingredient k in recipe i, and $\overline{g_k}$ denotes the average of $g_{k(i)}$.

We calculate the weights using the deviation and the distribution of the quantity of each ingredients. The ingredient that the deviation score is 50, ranks in the top 50 % of all ingredients(Fig.5). The ingredient that the deviation score is 60, ranks in the top 10 % of all ingredients. The ingredient that the deviation score is 80, ranks in the top 0.15 % of all ingredients. We divide up the weights W_k between 0 to 2, according to the deviation score. The score of cooking recipes are defined as shown in Eq.(9).

$$Score(R) = \sum_{k \in R} (I_k \cdot W_k) \quad (9)$$

W_k denotes the weights of the ingredient k in the target recipe. For example, if the deviation score is 50, the weights of the ingredient is 1.0. And if the deviation score is 60, the weights of the ingredient is 1.68.

III. Evaluate the Elicitation Accuracy of User's Preference and the Quantity of Each Ingredients

A. Experimental Condition

In order to assess the appropriateness of the scoring method, we conducted simple experiments. We used 8050 recipes extracted from a popular recipe search website in

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

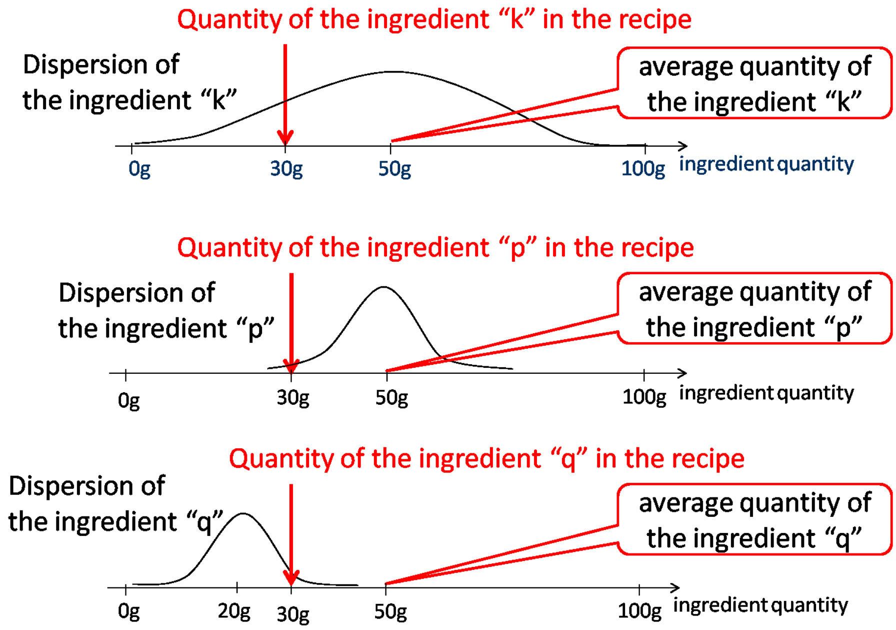

Fig. 4. The basic idea of the scoring method using the average and dispersion quantity of each ingredients

TABLE I
RANKING RESULTS OF 4 METHODS

1st	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
Recipe A	1	2	1	2
Recipe B	2	1	5	5
Recipe C	3	4	2	3
Recipe D	4	5	3	4
Recipe E	5	3	4	1

2nd	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
Recipe F	1	5	2	1
Recipe G	2	4	1	2
Recipe H	3	3	3	3
Recipe I	4	2	5	5
Recipe J	5	1	4	4

3rd	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
Recipe K	1	1	2	1
Recipe L	2	2	3	2
Recipe M	3	5	1	3
Recipe N	4	4	5	5
Recipe O	5	3	4	4

4th	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
Recipe P	1	1	2	1
Recipe Q	2	2	3	2
Recipe R	3	5	1	3
Recipe S	4	4	5	5
Recipe T	5	3	4	4

5th	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
Recipe U	1	2	1	1
Recipe V	2	3	5	5
Recipe W	3	2	2	4
Recipe X	4	1	3	2
Recipe Y	5	5	4	3

TABLE II
RANKING RESULTS BY NDCG VALUE

	1) correct ranking	2) ranking by popular website	3) browsing and cooking history	4) proposal method
nDCG	1	0.8996	0.9072	0.9381

Japan[8]. We used randomly selected recipes categorized as main dish.

We conducted experiment as follows.

1) We presented list of 5 recipes to subjects.

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

2) He/She ranks recipes manually according to the feelings that he would like to eat.
3) Repeat this sequence(1 and 2) 5 times.

Furthermore, beginning of this experiments, he/she inputs his/her cooking history of 30 days. Our system calculates the user's food preferences using this these history.

We compared 4 kinds of rankings, 1) ranking by user's subjective view(correct ranking), 2) ranking by the popular recipe search website, 3) ranking by user's food preferences(browsing and cooking history), and 4) ranking by proposal method(user's food preference and the quantity of each ingredients).

B. Evaluation Results

We evaluated the accuracy of reflect the user's preferences using user's food preferences(browsing and cooking history) and the quantity of each ingredients. The results are shown in TableI, and in addition, the nDCG value are shown in TableII. nDCG is a value that indicates a similarity to the correct ranking. In this evaluation, 1) ranking by user's subjective view is regarded as a correct ranking. Namely, nDCG value of 1) ranking by user's subjective view(correct ranking) is 1.0.

As shown in TableII, the result indicates that 4) proposal method got the highest nDCG value except 1) correct ranking and method of 3) browsing and cooking history got the second. It was found from the result that our methods provide better rankings than a conventional method though nDCG value of 2) ranking by popular website is still high enough. Thus, it seems reasonable to conclude that considering user' s food preference and the quantity of each ingredients in the recipe is very effective in order to achieve a personalized recipe recommender system.

According to impressions from the subjects, they sometimes considered the category of recipe and combination of ingredients. Therefore, we will try to consider the aspects for calculating recipe score in our future work.

IV. Conclusion

In this paper, we presented a scoring method for cooking recipe recommendation using the user's food preference and the quantity of each ingredients in the recipe. Our method estimates a user's food preferences from his/her past actions, such as through their recipe browsing and cooking history. We have proposed a recipe recommendation method based on the user's food preferences, that breaks down into their ingredients and scores them on the basis of the frequency of use and specificity of the ingredients. However, we consider that our previous method could not reflect user's preferences completely, so we propose a new recipe recommendation method based on user's food preferences and the quantity of the ingredients. Since our method can estimate the preferences through their browsing and cooking history, the user convey his/her preferences to the system without having to carry out any particular operation.

In order to verify the accuracy of reflecting the user's preferences using user's food preferences and the quantity of each ingredients, we conducted simple experiments. We compare 4 kinds of ranking results, 1) correct ranking , 2) ranking by the popular website, 3) ranking by the user's food preferences(browsing and cooking history), and 4) ranking by the proposal method(user's food preferences and the quantity of each ingredients). The experimental results by $nDCG$ values show that ranking results calculated by 4th method(using the user's food preferences and the quantity of each ingredients) is 0.9381, and it is similar to the user's subjective view.

References

[1] Cookpad. http://cookpad.com/, (Accessed 8 January 2014)
[2] Rakuten Recipe. http://recipe.rakuten.co.jp/ (Accessed 8 January 2014)
[3] Outline of Service (Cookpad Inc. Corporate Information), https://info.cookpad.com/outline_of_service/ (Accessed 8 January 2014)
[4] Yoko Mino and Ichiro Kobayashi, "Recipe Recommendation for a Diet Considering a User's Schedule and the Balance of Nourishment," *Proceedings of IEEE International Conference on Intelligent Computing and Intelligent Systems 2009*, pp.383-391, 2009.
[5] Shihono Karikome and Atsushi Fujii, "A System for Supporting Dietary Habits: Planning Menus and Visualizing Nutritional Intake Balance," *Proceedings of the 4th International Conference on Ubiquitous Information Management and Communication*, pp.386-391, 2009.
[6] Jill Freyne and Shalomo Berkovsky, "Intelligent Food Planning: Personalized Recipe Recommendation," *Proceedings of the 14th International Conference on Intelligent User Interfaces*, 2010.
[7] Mayumi Ueda, Mari Takahata and Shinsuke Nakajima, "Recipe Recommendation Method Based on User's Food Preferences", *Proceedings of the IADIS International Conference on e-Society 2011*, pp.591-594, 2011.
[8] Ajinomoto Park. http://park.ajinomoto.co.jp/ (Accessed 8 January 2014)

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)