Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 60 (2015) 986 – 993
Procedia Computer Science

19th International Conference on Knowledge based and Intelligent Information and Engineering Systems

Alternative-ingredient Recommendation Based on Co-occurrence Relation on Recipe Database

Ryosuke Yamanishi[a,*], Naoki Shino[a], Yoko Nishihara[a], Junichi Fukumoto[a,*], Aya Kaizaki[b]

[a]*Collage of Information Science and Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu 525-8577, Japan*
[b]*Ritsumeikan Global Innovation Research Organization, 1-1-1 Nojihigashi, Kusatsu 525-8577, Japan*

Abstract

This paper proposes a recommendation method of alternative-ingredients based on co-occurrence relation on recipe database. Currently, dishes are often cooked with reference to recipes on Website. Convenience to access so many and varied recipes encourages beginners to cook. Recipe on Website list ingredients used for a dish. However, for some reason, some of the listed ingredients cannot be used for the cooking; this paper defines such ingredient as "exchange-ingredient." To cook a dish, it should alternate exchange-ingredient and another one (i.e., alternative-ingredient). This paper proposes two algorithms to recommend alternative-ingredient. Through the cooking and tasting experiments, it was confirmed that the each of the proposed methods were effective for each intended purpose.

Peer-review under responsibility of KES International
Keywords: Cooking; Data mining; Bayesian method;

1. Introduction

We currently cook dishes with reference to recipes on varied types of information source such as book and Website. Though many books for cooking have been published, most people recently uses recipes on Website such as social network service and bulletin board systems. Cooking beginners are encouraged by the convenience to access so many and varied recipes.

A recipe on Website shows the ingredients that are needed for a dish and the procedure of the cooking. However, for some reason, some of the listed ingredients cannot be used for the cooking; foreign unique ingredients are difficult to prepare and the cause of allergy cannot be used. This paper defines such ingredient that cannot be used in the cooking as "exchange-ingredient." It is difficult to search recipes without exchange-ingredient on Web. Then, a cooking expert might use an alternative-ingredient instead of the exchange-ingredient. It is hard for cooking beginner determine the alternative-ingredient, because he/she has to consider at least the following matters to determine the alternative-ingredient;

* R. Yamanishi and J. Fukumoto. Tel.: +81-77-561-3416.
E-mail address: {ryama, fukumoto}@media.ritsumei.ac.jp

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2015.08.138

- Compatibility among ingredients used in a recipes should be good in cooking; not only the exchange-ingredient but also other ingredients in a recipe should be focused on.
- An ingredient whose role in a dish is similar to the exchange-ingredient should be prepared as the alternative-ingredient; on the other hand, the ingredient similar to the exchange-ingredient is objectionable because of allergic.

Especially, it would be difficult for cooking beginner to consider compatibility among ingredients.

This paper describes not recommendation problem but problem-solving for cooking. The goal of this study is recommendation of alternative-ingredient for cooking beginners. This paper, as the first step towards the goal, proposes a recommendation method of alternative-ingredients based on co-occurrence relation on recipe database. Co-occurrence relation means the relative frequency that more than two ingredients are concurrently used in a recipe. In this paper, two types of recommendation algorithm are compared; the one focuses on the similarity between exchange-ingredient and alternative-ingredient, and the other focuses on the role of exchange-ingredient in a recipe. This paper uses Japanese recipes data, however, the method does not depend on language and can be applied to other language.

2. Related Work

There is a lot of recipe data on Web such as "*Cookpad*[i]." Li *et al.* proposes RecipeCrawler that efficiently collects recipe data from Web[1]. His system collects contents of recipe on the Web with interactive annotation and description patterns.

Some researches trying to analyze recipes have been reported. Seki *et al.* propose a method to detect practical recipes for a dish from Website[2]. They mathematically modeled recipe, e.g., description of the cooking time, the recipe is written in detail, the number of steps, and so on. Their proposed method estimates whether a recipe is practical or not by using surport vector machine with the modeled features. Wang *et al.* propose a system to rewrite steps in a recipe to make it easy to understand[3]. In their system, steps in a recipe are fractionized, and a sentence representing multiple steps are itemized for each step.

Foodpairing, which is a method for identifying that ingredients are well together, has been proposed[4]. Foodpairing focuses on flavor of ingredients; if a flavor of an ingredient is similar to another's one, the ingredients are suitable to be used together. This theory applicables to western foods, however, not to asian foods because asian foods are too spicy and have much flavors. Chef Watoson[5] recommends additional ingredients for a dish based on analysis of recipe database. From the recommendation, professional chef can think up new way to cook the ingredients but regular people cannot. Blansché *et al.* propose a method to recommend alternative-ingredient for a dish[6]. Their method bases on the knowledge about what ingredient becomes alternative-ingredient of an ingredient, which is stored and edited with ontology. However, it is hard to cover varied ingredients based on static knowledge.

The proposed method in this paper uses recipes database on Web. Information of ingredients are extracted while excepting seasoners by using description patterns. And, alternative-ingredients are recommended based on the relationships among ingredients in recipes, without any other knowledge such as ontology.

3. The Proposed Method

This paper proposes two types of algorithms to recommend alternative-ingredient, and compares effectiveness of the each algorithm: recommendation based on compatibility between alternative-ingredient and ingredients in a recipe excepting exchange-ingredient, and recommendation based on similarity between exchange-ingredient and alternative-ingredient. Co-occurrence relations on recipe database is used in the both algorithms.

Fig. 1. Examples of a recipe on "*Kyou no Ryori*".

3.1. Recipe database

In this paper, "*Kyou no Ryori*[ii]"(translated to English:"A dish on Today") is referred as recipe database, where 18,650 recipes are stored. Recipes in the Website are proposed by professional chefs in TV program: "*Kyou no Ryori*" by NHK. The recipes are made for cooking in home, thus there are few unusual recipes in the Website. Then, the recipes are not classified into some categories: western, asian, Japanese, French, and so on.

The recipe page of "*Kyou no Ryori*" mainly consists of *Name*, *ingredients and its quantity*, and *Procedure* as shown as Fig. 1. This paper uses *ingredients and its quantity* for recommendation of alternative-ingredient. Both ingredients and seasoner are shown in *Ingredients and its quantity* without any annotations, thus it should be necessary to detect seasoners from the *ingredients*. Therefore, dictionary for seasoners is prepared. In the dictionary, ingredients whose quantities are "reasonable quantity" or "proper quantity" are recorded as entries; such ingredients are often seasoners through subjective analysis of recipes. Referring the dictionary, only ingredients are extracted for each recipe. In the database, 4,508 recipes, which are randomly selected from the Website, are stored. Also, the list of all kinds of ingredients in the database are prepared.

3.2. Co-occurrence relations among ingredients

The jth ingredient in a recipe database excepting seasoners is defined as Ing_j. Then, the information of co-occurrence relation $CR(Ing_j)$ is shown as the following;

$$CR(Ing_j) = (Fq(Ing_j, Ing_1), Fq(Ing_j, Ing_2), \ldots, Fq(Ing_j, Ing_k), \ldots, Fq(Ing_j, Ing_S)), \tag{1}$$

where, k shows also the index of the ingredient in the recipe database, $Fq(Ing_j, Ing_k)$ shows the relative frequency of recipes in which jth and kth ingredients are used together in the recipe database, and S shows the number of kinds of all ingredients in the database. The higher $Fq(Ing_j, Ing_k)$ means that the combination of jth and kth ingredients are often used in cooking. Note, $Fq(Ing_j, Ing_k)$ = NULL if $j = k$. And the set of the information of co-occurrence relation: $\boldsymbol{CR}$ is shown as the follow;

---

[i] http://cookpad.com/

[ii] http://www.kyounoryouri.jp/

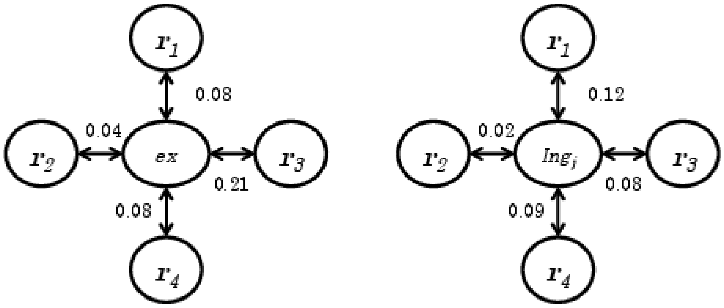

Fig. 2. The image example of the role of ex^i and Ing_j.

$$\boldsymbol{CR} = \{CR(Ing_1), CR(Ing_2), \ldots, CR(Ing_S)\}. \tag{2}$$

In this paper, $\boldsymbol{CR}$ is previously prepared from Website detailed in section 3.1. On the assumption that compatible ingredients are often used in a recipe together, alternative ingredients are recommended referring $\boldsymbol{CR}$.

3.3. Recommendation algorithms

In this paper, two types of algorithm are prepared and compared. Algorithm 1 focuses on compatibility between the alternative-ingredient and other ingredients in the recipes, which will be detailed in section 3.3.1. In this algorithm, similarity between the alternative-ingredient and the exchange-ingredient is not taken in the consideration. On the other hand, algorithm 2 focuses on similarity of the role in a dish between the ingredient and the exchange-ingredient, which will be detailed in section 3.3.2. This algorithm is envisioned for unavailability of an ingredient in a dish. The each algorithm will be detailed in the following subsections, respectively. Each ex^i and $r^i_p \in \boldsymbol{R}^i$ respectively shows an exchange-ingredient in ith recipe in the database and the pth remained ingredient, where $\boldsymbol{R}$ shows the set of remained ingredients.

3.3.1. Algorithm 1: Recommendation based on compatibility between alternative-ingredient and ingredients in a recipe excepting exchange-ingredient

The algorithm 1 recommends alternative-ingredient using Naive-Bayes filterring. The likelihood of $(Ing_j | Ing_j \neq ex^i)$ as the alternative-ingredient candidate for ith recipe: $Likelihood^i(Ing_j)$ is shown as the following equation;

$$Likelihood^i(Ing_j) = Fq(Ing_j) \times \prod_{p}^{N_i} Fq(r^i_p | Ing_j), \tag{3}$$

where, $Fq(r^i_p | Ing_j)$ is easily calculated with $\boldsymbol{CR}$ based on Bayes' theorem, and N_i shows the number of kinds of ingredient excepting seasoners in ith recipe.

The higher $Likelihood^i(Ing_j)$ means that the Ing_j is more compatible for $\boldsymbol{R}^i$. Based on this algorithm, indegrents that have higher *Likelihood* is recommended as alternative-ingredients. In this algorithm, the alternative-ingredient is recommended by using $r^i_p \in \boldsymbol{R}^i$. That is to say, the exchange-ingredient is not taken in the consideration.

3.3.2. Algorithm 2: Recommendation based on similarity between exchange-ingredient and alternative-ingredient

The algorithm 2 focuses on similarity between the role of exchange-ingredient and the one of alternative-ingredient. In this paper, the role of the exchange-ingredient in a recipe is modeled using the mutual information[7] between exchange-ingredient and the each remained ingredient. Fig. 2 shows the image example of the role of ex^i and Ing_j. In Fig. 2, the each value around an arrow shows the each mutual information between the connected nodes, which is previously calculated.

Table 1. Ingredient Information of “Chikuzen-ni” and “Green curry risotto,”

Title	The Exchange-ingredient	The remained ingredients
Chikuzen-ni	Burdock	Chicken, Taro, Carrot, LotusPea, Konjac
Green curry risotto	Grape Tomato	Rice, Chicken, Lyophyllum, Zucchini, Cheese for pizza

Table 2. Recommendation results of alternative-ingredient instead of “Burdock” for “Chikuzen-ni” with the each algorithm

The algorithm 1.

The Recommended alternative-ingredients	
1	Trefoil
2	Dried prawns
3	Pork
4	Radish
5	Leaf bud
6	Clam
7	Ginger
8	Shrimp

The algorithm 2.

The Recommended alternative-ingredients	
1	Radish
2	Carrot
3	Seaweed
4	Pumpkin
5	Pasta
6	Trefoil
7	Beef
8	Chinese Lemon

As the following equation, the role of ex^i: $role(ex^i)$ is shown as N_i-tuple vector;

$$role(ex^i) = (Mutual(ex^i : r^i_1), \ldots, Mutual(ex^i : r^i_q), \ldots, Mutual(ex^i : r^i_{N_i})), \quad (4)$$

where, $Mutual(ex^i : r^i_q)$ shows the mutual information between ex^i and r^i_q. Also, for jth ingredient, $role(Ing_j)$ is represented as ex^i is replaced to Ing_j in the equation (4). For example, in Fig. 2, the $N = 4$ and the $role(ex)$ can be shown as 4-tuple vector. Accordingly, the role of ex and Ing_j can be represented as $role(ex) = (0.08, 0, 04, 0.21, 0.08)$ and $role(Ing_j) = (0.12, 0, 02, 0.08, 0.09)$, respectively

The mutual information between jth and kth ingredients can be shown as the follows;

$$Mutual(Ing_j : Ing_k) = \sum_j \sum_k Fq(Ing_j, Ing_k) \log \frac{Fq(Ing_j, Ing_k)}{Fq(Ing_j)\, Fq(Ing_k)}, \quad (5)$$

where, $Fq(Ing_j, Ing_k)$ is assigned as referring $\boldsymbol{CR}$. The higher $Mutual(Ing_j : Ing_k)$ means that Ing_j and Ing_k are strongly related each other in the recipe database.

The Hamming distance between $role(ex^i)$ and $role(Ing_j)$: $HD(ex^i, Ing_j)$ is calculated for ith recipe. For example, $HD(ex^i, Ing_j) = 0.20$ in Fig. 2. This paper assumes that the lower $HD(ex^i, Ing_j)$ shows that the role of Ing_j is similar to the one of ex^i. This algorithm recommends ingredients that have lower $HD(ex^i, Ing_j)$ as alternative-ingredients for ith recipe.

4. Experiments to Recommend Alternative-ingredient

In order to verify the effectiveness of the proposed method, the experiments to recommend alternative-ingredients was conducted. Using the recommended ingredients, one of the authors cooked the dishes in which the exchange-ingredient was replaced to alternative-ingredients. Tastes and suitability of the alternative-ingredients for the exchange-ingredients were subjectively evaluated. The two types of algorithm were compared with each other for the effectiveness of the recommendation.

4.1. Alternative-ingredient recommendation

Two recipes selected from “*Kyou no Ryori*” were used for the experiments: “Chikuzen-ni” and “Green curry risotto.” The two recipes were different in genre; “Chikuzen-ni” is a traditional Japanese dish, in which chicken, root vegetables are boiled together and seasoned with soy-source, and “Green curry risotto” is original dish mixing Indian

Table 3. Recommendation results of alternative-ingredient instead of "Grape Tomato" for "Green curry risotto" with the each algorithm

The algorithm 1.

The Recommended alternative-ingredients	
1	Long green onion
2	Carrot
3	Ginkgo nuts
4	Bread
5	Citrus sudachi
6	Apple
7	Perilla
8	Juniper berry

The algorithm 2.

The Recommended alternative-ingredients	
1	Turnip
2	Parsley
3	Eryngii
4	Carrot
5	Cheese
6	Egg
7	Salted cod roe
8	Asparagus

Fig. 3. "Chikuzenn-ni" with "Pasta" instead of "Burdock."

and Italian. Table. 1 shows the exchange-ingredient and the remained ingredients in the experiment for "Chikuzen-ni" and "Green curry risotto." The exchange-ingredients were determined because we considered that the likes and dislikes depended on individual for these ingredients.

Table. 2 and Table. 3 each shows the results of the recommended alternative-ingredient for "Chikuzen-ni" and "Green curry risotto," respectively; instead of "Burdock" for "Chikuzen-ni," and "Grape Tomato" for "Green curry risotto." The results are sorted based on the likelihood of the ingredient for each algorithm. As shown in the tables, varied types of ingredients were recommended. It seemed that the results were caused by using "*Kyou no Ryori*" as the database, in which dishes in home were stored. Thus, varied co-occurrence relation among ingredients were obtained and it might cause the recommendation results as shown in Table. 2 and Table. 3. The recommendation does not consider the type of exchange-ingredient, thus meat and seafood were recommended as the alternative-ingredient of vegetable, e.g., instead of "Burdock," "Pork" and "Clam" were recommended. Consideration of types of ingredients and categorizing the recommended results would be our future work.

4.2. Cooking and subjective evaluation

One of the authors, who cooks twice in a week, cooked each "Chikuzenn-ni" and "Green curry risotto." For the demonstration, some ingredients were subjectively selected as alternative-ingredients for the exchange-ingredient. The ingredients used in the demonstration were as the follows;

- "Chikuzenn-ni"
  Algorithm 1: Trefoil, Dried prawns, Pork
  Algorithm 2: Seaweed, Pasta
- "Green curry risotto"
  Algorithm 1: Long green onion, Carrot, Bread, Apple
  Algorithm 2: Turnip, Parsley, Eryngii

Using the each alternative-ingredient instead of exchange-ingredient, "Chikuzenn-ni" and "Green curry risotto" were cooked. The process of cooking was conformed to the original recipe for each dish. Fig. 3 shows an example

Table 4. The results of the subjective evaluation.

Dish (Exchange-ingredient)	Algorithm	Alternative-ingredient	The evaluation A	B	C	D	E
"Chikuzenn-ni" (Burdock)	Algorithm 1	Trefoil	0	3	1	1	0
		Dried prawns	1	2	2	0	0
		Pork	0	5	0	0	0
	Algorithm 2	Seaweed	0	2	2	1	0
		Pasta	1	2	2	0	0
"Green curry risotto" (Grape Tomato)	Algorithm 1	Long green onion	0	3	2	0	0
		Carrot	0	3	2	0	0
		Bread	0	2	3	0	0
		Apple	0	0	3	2	0
	Algorithm 2	Turnip	0	3	1	1	0
		Parsley	1	1	3	0	0
		Eryngii	0	3	2	0	0

Table 5. The results of the subjective evaluation for each algorithm.

Algorithm	A	B	C	D	E
Algorithm 1	1(2.9%)	18(51.4%)	13(37.1%)	3(8.6%)	0(0%)
Algorithm 2	2(8.0%)	11(44.0%)	10 (40.0%)	2(8.0%)	0(0%)

of "Chikuzenn-ni" with "Pasta" instead of "Burdock." Five subjects ate the dishes cooked with each alternative-ingredient and evaluated the dishes with the following five options;

A Feeling for the alternative-ingredient approaches the one for the exchange-ingredient. The quality of the dish is on the caliber of the original one.

B Feeling for the alternative-ingredient is far from the one for the exchange-ingredient. However, the quality of the dish is either equaling or surpassing the original one.

C The taste of the dish is inferior as compared with the original one. That said, it is enough to eat and the dish is acceptable.

D It can be eaten, but it is not delicious.

E It tastes very terrible.

Table. 4 shows the results of the subjective evaluation. From the table, there was no answer for "E." Most subjects selected "B" or "C" for each alternative-ingredient. It was suggested that strange alternative-ingredients for the dish were not recommended by the both proposed algorithms.

Table. 5 shows the results of the subjective evaluation for each algorithm. As comparing algorithm 1 with 2, the difference of the evaluation between each algorithm was not founded. It was confirmed that both algorithms recommended acceptable alternative-ingredient for 90% as regarding "A," "B" and "C" were acceptable. Though algorithm 1 aimed to recommend ingredients similar to exchange-ingredient, answer "A," which means that alternative-ingredient approaches exchange-ingredient, is only 8% that is not so high. This result let us to reconsider the modeling of the role of the ingredient in a dish.

5. Conclusions and Future Work

This paper proposed alternative-ingredient recommendation based on co-occurrence relation on recipe database. Two types of recommendation algorithm were proposed for each purpose to exchange ingredient in a recipe. The one algorithm was based on Naive-Bayes filterring, which focused on the relation between alternative-ingredient and the remained ingredients. The other algorithm focused on similarity between the role of the exchange-ingredient and alternative-ingredient in a dish. Then, the role of ingredient was well modeled using mutual information between a given ingredient and the other ingredients in the recipe. The proposed recommendation algorithms should be compared with other recommendation algorithms such as collaborative filtering, in our future.

Through cooking demonstration with the recommended alternative-ingredient and subjective evaluation experiments, it was confirmed that the both algorithm recommended acceptable ingredients for over 90%. However, algorithm 1 might not recommend alternative-ingredient similar to the exchange-ingredient. This should be our future work, and we will reconsider the modeling of the role of ingredients using some information in a recipe data, e.g., process and quantity; the existing researches[8,9] would be useful to analyse process of cooking. In the future, as taking allergic information and nutrient factors in the consideration, the alternative-ingredient recommendation system for hospital diet will be developed based on the revealed findings in this paper: co-occurrence relation on recipe database can be applied to recommend alternative-ingredient. Also, accessibility of alternative-ingredient would be covered in the recommendation system for cooking in home.

Acknowledgements

This work was supported in part by Artificial Intelligence Research Promotion Foundation. We are appreciative of the discussions with Prof. Ebi at College of Sport & Health Science, Ritsumeikan University.

References

1. Li, Y., Meng, X., Wang, L., Li, Q.. Recipecrawler: Collecting recipe data from www incrementally. In: *Proc. of the International Conference Web-Age Information Manegement*. 2006, p. 263–274.
2. Seki, Y., Kouta, O.. Discriminating practical recipes based on content characteristics in popular social recipes. In: *Proc. of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2014, p. 487–496.
3. Wang, L., Li, Q.. A personalized recipe database system with usercentered adaptation and tutoring support. In: *Proc. of the Workshop on Innovative Database Research 2007*. 2007, p. 21–26.
4. Blumenthal, H.. Weird but wonderful - life and style. 2012. `http://www.theguardian.com/lifeandstyle/2002/may/04/foodanddrink.shopping/` (Retrieved on June 1, 2015).
5. IBM, . Cognitive cooking. 2014. `http://www.ibm.com/smarterplanet/us/en/cognitivecooking/` (Retrieved on June 1, 2015).
6. Blansché, A., Cojan, J., Dufour-Lussier, V., Lieber, J., Molli, P., Nauer, E., et al. Taaable 3: Adaptation of ingredient quantities and of textual preparations. In: *Proc. of the 18th International Conference on Case-Based Reasoning*. 2010, p. 189–198.
7. Church, K.W., Hanks, P.. Word association norms, mutual information, and lexicography. In: *Proc. of the 27th Annual Meeting of the Association for Computational Linguistics*. 1989, p. 22–29.
8. Hamada, R., Okabe, J., Ide, I., Satoh, S., Sakai, S., Tanaka, H.. Cooking navi: Assistant for daily cooking in kitchen. In: *Proc. of the 13th annual ACM international conference on Multimedia*. 2005, p. 371–374.
9. Buykx, L., Petrie, H.. What cooks need from multimedia and textually enhanced recipes. In: *Proc. of 2011 IEEE International Symposium on Multimedia*. 2011, p. 387–392.