Alternative-Ingredient Recommendation Based on Correlation Weight for Thai Recipes

1st Wanvimol Nadee
Information Technology Division, Faculty of Science
Maejo University
Chiangmai, Thailand
ORCID 0000-0001-5577-4108

2nd Sayan Unankard
Information Technology Division, Faculty of Science
Maejo University
Chiangmai, Thailand
ORCID 0000-0002-8443-5356

***Abstract*—Cooking is a popular activity. Most people recently use recipes on the Website, which show the ingredients that are needed for a dish and the method of cooking. However, some of the listed ingredients sometimes are not able to prepare or use for cooking for some reasons: allergies, personal preference, or some ingredients are not available in the kitchen, etc. Therefore, it should recommend the alternative-ingredients that are similar to an exchange-ingredient and suitable with remained ingredients in the recipe. This paper proposes a recommendation system of alternative ingredients for Thai recipes. The recommended method is based on the combination of the smoothed correlation weight function and graph-based approach on the Thai recipe database. Our research contributes to enhancing the cooking beginners to cook Thai food and prepare ingredients conveniently by using the alternative-ingredients.**

***Keywords*—Alternative ingredient recommendation, Food innovation, Text mining, Data mining, Recommendation system, Thai recipes.**

I. Introduction

The recipe websites are a popular source of sharing recipes and cooking methods from different cuisines. They open a new door for people who are looking for their favorite dishes and want to try a new dish that they would never try. As the recipe websites make recipes easier to access, thereby people turn to recipe websites where they can share and find information easily. A recipe on the Website shows the ingredients that are needed for a dish and the cooking process. However, sometimes people want to cook an unfamiliar dish and they could not get all the ingredients listed on the recipe [1]. In other words, some of the listed ingredients sometimes are not able to prepare or use for cooking for some reasons: allergies, religions of individuals living, personal preference, some ingredients are not available in the kitchen or specific ingredients are not able to source the right ones at the moment [2]. To address these problems, we want to discover a suitable alternative-ingredient that goes well with the remained ingredients in that recipe for substitution.

Search and recommendation systems play an important role in the way people choose what they eat [3]. There has been much research on recipe recommendation systems that were developed for different purposes. Most existing works mainly focus on the analysis of recipes based on recipe content (e.g., ingredient). For example, Sajadmanesh et al. [4] used the ingredients, taste, and cuisine information to understand culinary habit around the world. Elsweiler et al. [3] exploited the online recipe to help people to choose a healthier meal. They investigated users' behavior of how people perceive and select recipes. The machine learning techniques are applied to find the replacement recipes and predict the preferred recipes. Moreover, Shino et al. [2] and Yamanish et al. [5] conducted the alternative ingredients recommendation systems based on co-occurrence relation among ingredient category information. The basic idea of choosing alternative ingredients is based on the similarity and compatibility of ingredients in the same category. They mentioned the ingredients belonging to the same category of the exchanged-ingredient were more suitable as alternative-ingredient. Whereas Lui et al. [1] used the ingredients co-occurrence frequency along with category importance calculated from recipe data for each replacement. Their method can recommend suitable ingredients with a diversity of ingredients.

This paper proposes a recommendation system of alternative ingredients for Thai recipes. Thai cuisine is one of the world's great cuisines and is also a unique food. Cooking Thai food might not be easy for beginners regarding preparing the ingredients listed in the recipe. Due to they do not have a specific or main ingredient listed in the refrigerator. Thus, we need to find another ingredient that is suitable with the remained ingredients and in the same category in that recipe for replacement. For instance, we want to make Thai green curry but we cannot find the chicken in our kitchen. We might desire to use something other to replace chicken such as pork or beef or meat or seafood. The goal of this study is to enhance the cooking beginners to cook Thai food and prepare ingredients conveniently by using the alternative-ingredients. Our basic idea is to exploit the concept of co-occurrence of ingredients along with the smoothed correlation weight function to calculate the important weight give to each ingredient on recipe data. Then we leverage the strong relations among exchange-ingredients and other ingredients using the graph-based approach to capture the alternative-ingredients. The challenge of this work is how to discover the alternative ingredients that are similar and proper to an exchange-ingredient in the complex relationships between

978-0-7381-3053-8/21/$31.00 ©2021 IEEE

ingredients.

The main contributions of this work are an approach based on the combination of three techniques that can find appropriate ingredients when recommending alternative ingredients. Furthermore, we use what we learn to establish the preliminary model of alternative ingredient recommendation for Thai recipes which considers only the main ingredient. The evaluation of our method based on the Thai recipe dataset collected from the Yummly website. The experimental results were compared with these two methods which demonstrate the proposed approach can improve the effectiveness of recommendations.

The rest of this paper as follows: Section 2 discusses related work in recipe recommendation systems and the alternative ingredient recommendation systems; Section 3 presents the proposed alternative ingredient recommendation method; Section 4 describes our experiments and analysis; Section 5 summarizes our findings and suggests possible future work.

II. Related Works

The study of food recommendation systems has been increasing due to their relevance for healthy living. The recipe content on the Web provides valuable information to the food research community. The studies on recipe mainly focus on recipe recommendation, menu recommendation, and ingredient replacement [1]. For example, food pairing focuses on a similar flavor of ingredients; if a flavor of an ingredient is similar to another one. The ingredients are suggested to be used together. The food pairing hypothesis has been studied for several modern cuisines. Earlier studies of food pairing are Ahn et al. [6] and Varshney et al. [7] examined a flavor network involving ingredients derived from recipe to capture their flavor bonds. Mao et al. [8] and GUO et al.[9] utilized the flavor pairing concept to recommend a set of dishes from the various Chinese regional cuisines for a certain flavor preference in terms of flavor similarity. The cosine similarity and the TF-IDF algorithm are applied to choose the dishes with the most similar flavors and recommend them to the users. Our work is related to an alternative ingredient recommendation system, which is used to recommend an alternative if a particular ingredient is not present or cannot be used in the recipe [10]. Previous works are proposed in [1], [2], [5], [11], the authors conducted the approaches based on co-occurrence relation among ingredient category information. The basic idea of choosing alternative ingredients is based on the similarity and compatibility of ingredients in the same category. They mentioned the ingredients belonging to the same category of the exchanged-ingredient were more suitable as alternative-ingredient. They define such an ingredient that cannot be used in the cooking as “exchange-ingredient”. Their method only recommending the ingredient in the same category, the suitable ingredients in different categories are not taken to consideration. While Lui et al. [1] mentioned that their algorithm ignores suitable ingredients in different categories. Then, they utilized the ingredients co-occurrence frequency along with category importance calculated from recipe data for each replacement. Their method provides suitable ingredients with a diversity of ingredients. Maheshwari and Chourey [10] proposed a machine learning model to innovate new dishes and to help people allergic to certain ingredients by recommending alternate ingredients. It has shown that the ingredient co-occurrence frequency is widely used in the alternative-ingredient recommendations. However, our work relies on smoothed correlation weight function to calculate the important weight based on the relation among ingredients in the recipe. A graph-based approach is utilized to discover the relation between the exchange-ingredient and the replace ingredients.

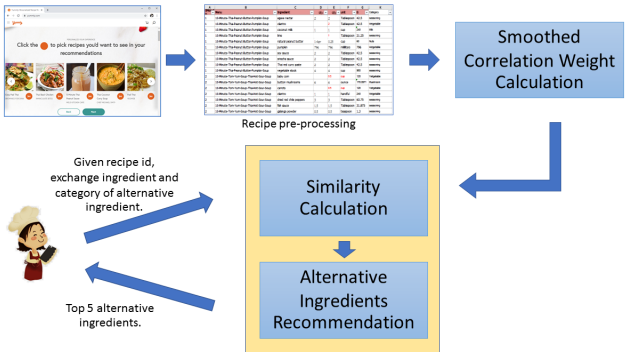

Fig. 1. The architecture of our system.

III. Proposed Method

In this task, we present our approach to recommending alternative ingredient for Thai recipes. Figure 1 shows a conceptual framework of our approach. It can be divided into three steps: Recipe pre-processing, Smoothed correlation weight calculation, Similarity calculation and Alternative ingredients recommendation. The details of each step of our approach are given in the following.

A. Recipe Pre-processing

In order to improve the quality of our dataset and the performance of the proposed approach, the pre-processing is performed. Thai recipes are extracted from the dataset. We manually assigned each ingredient to a category and renamed the ingredient to the proper one such as “boneless beef steak”, “beef steak” and “beefsteak” to “beef steak”. Example of the dataset after pre-processing is shown in Table I.

TABLE I
Example of the dataset after pre-processing.

Recipe Name	Ingredient	Quantity	Unit	Category
Tom Kha Gai	chicken	1	pound	meat
Tom Kha Gai	coconut milk	2	can	milk
Tom Kha Gai	coconut oil	3	tablespoon	oil
Tom Kha Gai	fish sauce	1	tablespoon	seasoning
Tom Kha Gai	galangal	2	tablespoon	vegetable
...	...	..	...	...

B. Smoothed correlation weight calculation

In this task, we aim to find the relationship between two ingredients. Our intuition is that ingredients frequently co-occur when they have meaningful relationships between them. To extract the set of co-occurring ingredients, the directed edge-weighted graph is created. The edge is created if the correlation weight score between two ingredients greater than zero. Example of the directed edge-weighted graph is show in Figure 2. We adopt the smoothed correlation weight function to calculate the semantic correlation weight between two ingredients. All ingredients are used to calculate the correlation weight between two ingredients because some ingredients can sometimes be both seasoning and primary ingredients in different recipes. The formula is shown below:

$$A = (n_{xy} + \frac{n_x}{N})/(n_x - n_{xy} + 1) \quad (1)$$

$$B = (n_x - n_{xy} + \frac{n_x}{N})/(N - n_x - n_y + n_{xy} + 1) \quad (2)$$

$$SCW_{x,y} = log(\frac{A}{B}) \quad (3)$$

where $SCW_{x,y}$ is a correlation weight score between ingredient x and y, n_x is a number of recipes containing ingredient x, n_y is a number of recipes containing ingredient y, n_{xy} is a number of recipes containing both ingredient x and y, while N is the total number of recipes.

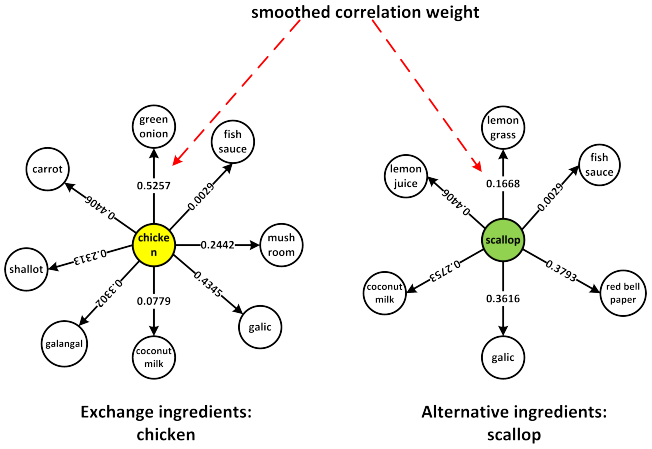

Fig. 2. Example of the directed edge-weighted graph.

C. Similarity calculation and Alternative ingredients recommendation

The problem that we address in this section is how to find the similarity of ingredients for the alternative recommendation. For a given category, a direct edge-weight graph of each ingredient in a given category is created. Each graph of candidate ingredient will be compared with the exchange ingredient graph. The cosine similarity between the two graphs will be used to find the similarity between the exchange ingredients and candidate alternative ingredients. The top 5 ingredients with the highest similarity scores will be recommended to the user. The cosine similarity function is shown below:

$$iSim(ex, al) = \frac{\sum_i (SCW_{ex_i} \times SCW_{al_i})}{\sqrt{\sum_j SCW_{ex_j}^2} \times \sqrt{\sum_k SCW_{al_k}^2}} \quad (4)$$

where $iSim(ex, al)$ is a similarity value between ingredient ex and al while ex and al are directed edge-weighted graphs of exchange ingredient and alternative ingredient respectively.

IV. Experiments and Evaluation

A. Dataset

The Yummly-28K recipe dataset collected from Yummly and provided by www.lherranz.otg/datasets/ to develop our algorithms and evaluate the performance. We collected only Thai recipes which consist of 832 dishes and 868 ingredients.

B. Baseline

In order to evaluate our approach, we compare our approach performance with the co-occurrence frequency approach adopted from [2]. We calculated the occurrence frequency of an ingredient and the co-occurrence frequency of two ingredients. For solving an orthographical variant with a cooking ontology, we cannot perform this step because we do not have a cooking ontology. Then, the compatibility score between the ingredient and the recipe is calculated. The top 5 alternative ingredients with the highest scores of the proportion of alternative ingredients with compatibility score of all ingredients in a recipe database will be recommended.

C. Evaluation

In this section, we evaluate the performance of our approach. The experiments were conducted with four recipes. The four recipes of our evaluation are shown in Table II. The top 5 ingredients for a selected category are recommended. The example of the top 5 recommended ingredients is shown in Table III. For preliminary evaluation, we asked 7 participants who have Thai cooking experience from 6 countries (Thailand, Australia, Korea, Norway, Germany, and the other two from the USA) to evaluate whether the recommended ingredients were suitable as an alternative or not based on their experience on cooking. The participant will give a score (1-5) for each recommended ingredient. More point means the ingredient is more suitable for the replacement. The result of our evaluation is shown in Table IV. However, for the taste evaluation, we will conduct it in our future work.

TABLE II
THE FOUR RECIPES OF OUR EVALUATION.

Menu Name	Exchange ingredient	Alternative Category
Tom Kha Gai	chicken	meat, seafood
Thai Green Curry	coconut milk	milk, nuts
Thai Basil Beef	beef	meat, seafood
Thai Papaya Salad	green papaya	meat, seafood, vegetable, fruit

In Table IV, our approach achieves a higher average score over the baseline. The results show that most participants

TABLE III
THE EXAMPLE OF TOP 5 RECOMMENDED INGREDIENTS TO REPLACE "GREEN PAPAYA" FOR THAI PAPAYA SALAD.

	Vegetable	**Fruit**	**Meat**	**Seafood**
Baseline	ginger (9) lemongrass (11) red bell pepper (11) green onion (10) shallot (11)	mango (29) pineapple (17) coconut (11) coconut flakes (8) avocado (15)	chicken breast (18) chicken (16) chicken thighs (16) sirloin steak (18) boneless chicken thighs (17)	shrimp (20) red snapper (13) salmon fillet (19) prawns (18) whitefish fillets (13)
Our approach	green long beans (31) salad (24) green cabbage (26) shallot (11) cucumber (33)	mango (28) avocado (15) coconut (11) pomegranate seeds (10) pineapple (17)	chicken thighs (17) pork (17) ground pork (16) sirloin steak (18) boneless pork lion roast (15)	shrimp (19) salmon fillet (18) whitefish fillets (13) halibut fillets (13) prawns (18)

(x) is a total score from 7 participants given for each recommended ingredient.

TABLE IV
THE AVERAGE SUITABLE SCORE (1 - 5) OF OUR EVALUATION.

Participant	**Country**	**Baseline**	**Our Approach**
1	Thailand	2.52	**2.84**
2	Australia	3.30	**3.56**
3	Korea	3.08	**3.40**
4	Norway	2.66	**2.70**
5	USA	2.78	**2.98**
6	USA	**2.46**	2.20
7	Germany	2.46	**2.64**
	Average	**2.75**	**2.90**

prefer our recommended results except one participant from the USA. In addition, participants from Australia and Korea gave average scores greater than 3.00 while participants from the USA and Germany gave average scores less than 2.50. This may be due to the variety of ingredients in each country and individual preferences.

V. CONCLUSIONS

In this paper, an alternative-ingredient recommendation system for Thai recipes is proposed. The goal of this study is to effectively recommend alternative ingredients to replace a specific ingredient in a recipe. Our contributions are summarized as follows: (1) we introduced a recommendation approach of alternative ingredients for Thai recipes and (2) the incorporation between smoothed correlation weight and graph-based approach achieved better results than the baseline. In future work, we will further consider the quantity of each ingredient in the recipe. Cooking demonstration and taste evaluation will be conducted.

ACKNOWLEDGMENT

This research was supported by The Coordinating Center for Thai Government Science and Technology Scholarship Students (CSTS), National Science and Technology Development Agency (NSTDA).

REFERENCES

[1] K.-H. Liu, H.-C. Chen, K.-T. Lai, Y.-Y. Wu, and C.-P. Wei, "Alternative ingredient recommendation: A co-occurrence and ingredient category importance based approach." in *PACIS*, 2018, p. 298.

[2] N. Shino, R. Yamanishi, and J. Fukumoto, "Recommendation system for alternative-ingredients based on co-occurrence relation on recipe database and the ingredient category," in *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*. IEEE, 2016, pp. 173–178.

[3] D. Elsweiler, C. Trattner, and M. Harvey, "Exploiting food choice biases for healthier recipe recommendation," in *Proceedings of the 40th international acm sigir conference on research and development in information retrieval*, 2017, pp. 575–584.

[4] S. Sajadmanesh, S. Jafarzadeh, S. A. Ossia, H. R. Rabiee, H. Haddadi, Y. Mejova, M. Musolesi, E. D. Cristofaro, and G. Stringhini, "Kissing cuisines: Exploring worldwide culinary habits on the web," in *Proceedings of the 26th international conference on world wide web companion*, 2017, pp. 1013–1021.

[5] R. Yamanishi, N. Shino, Y. Nishihara, J. Fukumoto, and A. Kaizaki, "Alternative-ingredient recommendation based on co-occurrence relation on recipe database," *Procedia Computer Science*, vol. 60, pp. 986–993, 2015.

[6] Y.-Y. Ahn, S. E. Ahnert, J. P. Bagrow, and A.-L. Barabási, "Flavor network and the principles of food pairing," *Scientific reports*, vol. 1, p. 196, 2011.

[7] K. R. Varshney, L. R. Varshney, J. Wang, and D. Myers, "Flavor pairing in medieval european cuisine: a study in cooking with dirty data," *arXiv preprint arXiv:1307.7982*, 2013.

[8] X. Mao, S. Yuan, W. Xu, and D. Wei, "Recipe recommendation considering the flavor of regional cuisines," in *2016 International Conference on Progress in Informatics and Computing (PIC)*. IEEE, 2016, pp. 32–36.

[9] L.-R. GUO, S.-Z. YUAN, X.-H. MAO, and Y.-N. GU, "A recipe recommendation system based on regional flavor similarity," *DEStech Transactions on Computer Science and Engineering*, no. cimns, 2017.

[10] S. Maheshwari and M. Chourey, "Recipe recommendation system using machine learning models," 2019.

[11] N. Shino, R. Yamanishi, and J. Fukumoto, ""co-occurrence relation" and "ingredient category" recommend alternative-ingredients," in *New Trends in E-service and Smart Computing*. Springer, 2018, pp. 1–18.