Complexity and Similarity of Recipes based on Entropy Measurement

Article *in* Indian Journal of Science and Technology · July 2016

DOI: 10.17485/ijst/2016/v9i26/97324

4 authors, including:

Yun-Jung Lee
Pusan National University
39 PUBLICATIONS **85** CITATIONS

SEE PROFILE

Hwan-Gue Cho
Pusan National University
262 PUBLICATIONS **1,494** CITATIONS

SEE PROFILE

Seong-Min Yoon
Pusan National University
165 PUBLICATIONS **1,987** CITATIONS

SEE PROFILE

Complexity and Similarity of Recipes Based on Entropy Measurement

Su-Do Kim[1], Yun-Jung Lee[2], Hwan-Gue Cho[3], Seong-Min Yoon[4*]
[1,2]Research Institute for Social Criticality, Pusan Nat'l Univ., Korea, {kimsd, leeyj01}@pusan.ac.kr
[3] Dept. of Computer Science, Pusan Nat'l Univ., Korea, hgcho@pusan.ac.kr
[4*] Dept. of Economics, Pusan Nat'l Univ., Korea, smyoon@pusan.ac.kr

Abstract— Foods and cuisines have become a significant cultural trend and various recipes are opened and shared on the internet. Foods are usually evaluated and determined by subjective factors such as the personal cooking ability or the personal preference. In this paper, we newly define the objective method for measuring the complexity and the similarity of the recipe using the probabilistic entropy measurement. Whether two recipes are similar or not is determined by the entropy measurement of common ingredients. And we construct the social network of recipes by the similarity degree. Through the recipe network, we can measure the distance of the cuisines and create the comprehensive concept map of recipes that observe the whole correlation between recipes.

Keywords-Recipe; Ingredient; Entropy measure; Recipe complexity; Similarity;

1. INTRODUCTION

Cooking is the fundamental and creative activity to satisfy the basic human instinct and people's common interest. Recently, foods and cuisines have become a new significant cultural trend in broadcasting and media. New kinds of dishes and recipes are being constantly created and developed by people. And they are published and shared on the internet.

In this paper, we explore how recipes are connected by social relationship, measured by some objective criteria, and shared by common elements in the abstract space of cooking. Various recipes, including both traditional cuisine and fusion cuisines which combine different culinary traditions, have been introduced by people on the web. But the recipe is closely related to the cognitive characteristics of a person who describes it. We can see the difficulty degree of the cooking preparation or the cooking procedure in the recipe information. It is determined to be very subjective factors based on the personal cooking ability, personal preference or cultural background. There is no scientific basis for the evaluation of the difficulty of the recipe.

In this paper, we provide the basis for objectively evaluating the complexity of a recipe to prepare and to cook. Our objective evaluation is based on the entropy concept used in the field of information engineering. And we measure the similarity between recipes by using common ingredient entropy. Finally, we construct the recipe similarity network that is connected by the similarity degree between recipes. The recipe network provides a mapping method that applies the correlation of recipes to the distance space of recipes and can be expressed as the comprehensive concept map for recipes.

Using this recipe map, we can observe at a glance the whole correlation of recipes and understand the issues as follows: Which food is similar to a particular food? How the foods are clustered? Is there any possibility one can create new cooking that combines some cuisines? Also, we can find in the objective viewpoint to what cooking will be made when different cuisines are blended and which path is the most reasonable in the ease of cooking preparation when we want to change any food to other foods. Through the recipe network, it can be found a rational and an objective manner in terms of the recipe preparation and the recipe procedure.

2. RELATED WORK

Several researches to find a similar recipe have been conducted. The main subject of the recent researches in recipe is about the connection of ingredients through the social network. The social network reveals a new insightful concept in various fields [1-3]. Wang et al. [4] analyzed the flow of cooking procedure and the order to add ingredients to cook. They proposed a method to find the repetitive cooking structure appeared in Chinese cuisine. They showed the similarity measurement based on the frequent patterns and the filtering algorithm to prune unrelated data. Then they represented the recipes as cooking graph using the similarity measurement. Teng et al. [5] showed the extensive study on Western food and ingredient. They constructed two types of networks, the complement network and the substitute network, to capture the relationship between ingredients. If two ingredients tend to co-occur frequently, two ingredients are connected to each other as they are complementary.

Using the complement network, they recommended the replaceable ingredient. They showed American regional characteristics in cuisine through the substitute network that derived from a user-generated suggestion for modifications. If any ingredient in a recipe is not easy to get, the substitute network provides the replaceable ingredient.

Some works which focus on the food recommendation system depends on personal nutrition, preference, or status. Usually, the system suggests user's food preferences based on their past recipe ratings or history. They reveal how the ingredient is recommended for the good food and how to replace a cooking material through analyzing the ingredients and the cooking verbs [6-8]. According to the study, they showed the alternative ingredients, 'Carrot' instead of 'Salami' and 'Orange' instead of 'Blue Pepper' in 'Pizza Toast' recipe.

Another research analyzed various factors of an ingredient and presented the substitute ingredient network by considering the nutrition and the chemical component [9-12]. In this paper, we consider dishes or ingredients as an independent object, construct the new relationship based on the correlation between objects, and find the characteristic by analyzing the social network [13].

3. DATA STATISTICS

3.1 Data collection and refinement

We collected 897 recipes from 'Allrecipes.kr', the Korean website of 'Allrecipes', for creating the recipe network. 'Allrecipes' operates 18 web sites around the world and it is one of the famous recipe sharing website which has about 1 billion users per year.

The recipe contains the information to cook food: ingredients, cooking time, and cooking procedure. Each ingredient is described with the ingredient name, the amount of the ingredient, and the ingredient shape. To extract only ingredient names from the recipe, we removed and refined the unnecessary data in the recipe stored in a text file. The final extracted names are 421. And we also extract the cooking verbs in the same way as the ingredient names because the cooking verb has the meaning of describing the cooking method or cooking procedure. The final extracted verbs are 110.

Table 1 shows the steps for extracting the ingredient names from the recipe. The number of country recipes and the extracted ingredient name is shown in Table 2.

3.2 Statistical analysis of ingredients

For analyzing what ingredients are used very often for cooking or the other way around, we calculated the frequency rate of the recipes and the ingredients. The frequency shows what ingredients are used mostly or not in the recipes. Table 3 shows the top 10 and the bottom 10 at the top of ingredient rankings and verb rankings. The most used ingredient in 897 recipes is 'Garlic' that is included in 484 recipes, about 54 percent of total recipes. The ingredients such as 'Salt,' 'Welsh Onion,' 'Soy Sauce,' are used mostly in both Korean foods and foreign foods. The most used cooking verb is 'Boil' that is used in 461 recipes, about 51 percent of total recipes.

The frequency distributions of the number of recipes and ingredients are shown in Figure 1 and 2. Figure 1 shows the frequency of the number of recipes related to the total number of ingredients in a recipe. Some recipes need so few ingredients for cooking, but some recipes need so many ingredients more than 20 ingredients. The average number of ingredients that are used in a recipe is 10. Figure 2 shows the frequency of the number of ingredients related to the number of recipes that include the same ingredient. The ingredient such as 'Garlic' is included in 484 recipes, but 'Licorice Root' is used only in 1 recipe. The average number of recipes that include any ingredient is about 21.

<insert figure 1 here>

<insert figure 2 here>

4. ENTROPY MEASURE AND RECIPY COMPLEXITY

4.1 Entropy measure

In this paper, we used the method of the entropy measure by the probability. Entropy is a quantitative measure of uncertainty and it is the expected value of information. If two ingredients, 'Garlic' and 'Licorice root,' are used in a recipe, the effects of two ingredients to food may be different. Generally speaking, it is accepted that two recipes are considered similar if they share a lot of common ingredients or cooking verbs. But it is difficult to say that two recipes are similar if they share common ingredients such as 'Garlic', 'Salt', or 'Welsh Onion'. Because 'Garlic' and 'Salt' are the most used ingredients in recipes and they have low information on the characteristic of cooking. But if two recipes include an ingredient such as 'Persimmon' or 'Licorice Root', the similarity of two recipes is increasing. Because 'Licorice Root' is rarely used in cooking, it has more information to calculate the

characteristic of cooking and the similarity between recipes than 'Garlic'. The ingredient such as 'Licorice Root' has higher entropy value than the ingredient such as 'Garlic'. We measure the ingredient entropy and the recipe entropy using the frequency probability.

First, we define R is the set of all recipes we collected and we assume that $|R| = m$. And r_i is the recipe that includes an ingredient i_x. We can define $P(i_x)$, the probability of an ingredient i_x.

$$P(i_x) \equiv P_x = \frac{|\{r_i \mid i_x \in r_i\}|}{|R|} \quad (1)$$

Using $P(i_x)$, we can define the entropy of an ingredient i_x, $E(i_x)$ as follows.

$$E(i_x) = -\log P(i_x) \quad (2)$$

We can know that the rarer the ingredient is, the higher $E(i_x)$ is. And by summing up all entropies of ingredients contained in a recipe, we also define $E(r_x)$, the entropy of r_x as follows.

$$E(r_x) = \sum_{i_k \in r_x} E(i_k) \quad (3)$$

4.2 Recipe complexity

The recipe entropy is an objective criterion to measure the complexity of cooking. If recipe entropy is low, it means that the ingredients of a recipe are commonplace or the number of ingredients is small. This recipe is easy to cook and has low complexity to prepare food. In contrast, high entropy means that the recipe needs rare or many ingredients. As the preparation of ingredients is not easy, the complexity of the recipe becomes higher.

We show the distribution of recipe entropy in Figure 3. The *X*-axis is the recipe entropy by the summing of ingredient entropies and y-axis is the recipe entropy by the summing of cooking verb entropies. As the *x* value increases, it means that the recipe includes more rare ingredients or the number of ingredients. The complexity of the recipe preparation is high. As the y value increases, the complexity of the recipe procedure is high.

<insert figure 3 here>

Table 4 and 5 show the lists of recipes and ingredients of each recipe according to the ranking of the recipe entropy. If you want to cook 'Boiled Potato', you just prepare only 3 ingredients, 'Potato', 'Salt' and 'Sugar'. But if you cook 'Braised Dongtae Seafood', you should prepare 27 ingredients and unusual ingredients such as 'Sweet Pumpkins' and 'Seafood'.

Table 6 shows high ranked recipes by the cooking verb entropy. The higher recipe entropy by verbs increases the complexity of the cooking procedure. The highest recipe is 'Jeolla-Province Mosi-Songpyeon'. To cook 'Jeolla-Province Mosi-Songpyeon', the type of the verbs is 18 and the total number of duplicated verbs is 28. This needs 28 actions to cook: boil, parboil, fry, soak, pound, etc. The lowest recipe is 'Healthy Silk-Fowl Soup' that does only 'Boil' to cook. Cooking verb is a word to explain the cooking procedure. If a recipe has the higher entropy, it is difficult to cook and has the higher complexity of the cooking.

5. Construction of Recipe Similarity Network

5.1 Recipe similarity

Recipe similarity is the objective information to distinguish whether the cuisines are similar or not. It is hard to measure the similarity between recipes by the number of common ingredients. For example, two recipes use a lot of common ingredients such as 'Garlic, 'Salt', and 'Soy Sauce'. Can we say that two recipes are similar? The number of common ingredients is not a sufficient measure, without considering the frequency of ingredients. We consider these factors. The entropy of ingredients with considering the rarity of ingredients would be a good measure for deciding the similarity of the recipe.

$\mathrm{Sim}(r_x, r_y)$ will be one universal similarity of two recipes $\mathrm{r_x}$ and $\mathrm{r_y}$ in the following.

$$\mathrm{Sim}(r_x, r_y) = \frac{\sum_{i_k \in \{r_x \cap r_y\}} E(i_k)}{\sum_{i_k \in \{r_x \cup r_y\}} E(i_k)},\ 0 \le Sim \le 1 \quad (4)$$

Table 8 shows the ranking of the similarities calculated by summing up the entropies of common ingredients between recipes. The Similarities between recipes, 'Candied Sweet Potato' and 'Sweet Potato Stick', 'Seasoned Bean Sprouts' and 'Cooked Bean Sprouts', 'Balloon Flower Salad' and 'Balloon Flower and Cucumber Salad', and 'Medicinal Rice' and 'Medicinal Rice by Electric Cooker' is 1. We know the recipes are practically the same cooking and the similarity by our entropy measure is nearly 1. The recipe title is written by the user when he registers a recipe on a website. In case of different recipe titles but the same cooking, we cannot distinguish by only recipe titles, whether two recipes are equal or not. The recipe similarity is the objective information to distinguish whether two recipes are similar or different. And it would be useful for searching similar cuisines.

5.2 Recipe similarity network

We can construct the recipe graph $G_R(V, E)$ by making edge (r_x, r_y) between two recipes r_x and r_y if $\mathrm{Sim}(r_x, r_y) \ge t_0$ for a threshold t_0 given. Figure 5 shows the Korean recipe network by $\mathrm{Sim}(r_x, r_y) \ge t_0 (t_0 = 0.36)$. As the edge is connected with the order of similarity from highest to lowest, similar recipes are grouped together. In figure 4, a large group composed of 302 recipes is displayed with red color in the center, small groups displayed around are composed of two or more recipes, and 328 recipes displayed in outer space are showing a very low correlation to each other. Figure 5 shows the connecting process to the network according to the degree of the similarity between the recipes.

<insert figure 4 here>

<insert figure 5 here>

Figure 6 shows the high related recipes in a Korean recipe network by $\mathrm{Sim}(r_x, r_y) \ge t_0 (t_0 = 0.43)$. For Example, if you cook 'Julienne Korean Balloon-Flower Salad' today, you will want to cook a different dish tomorrow, not same food. However, if you plan to cook a new dish with different ingredients of not existing, the cooking will be difficult because of buying new ingredients. It can be easy to prepare the dish with the existing ingredients as possible. The complexity of the recipe preparation will be low if we can cook a different cuisine with existing ingredients. By changing a few ingredients, 'Balloon-Flower Salad' can be easily changed to 'Cucumber Salad' or 'Turban-Shell Salad'.

<insert figure 6 here>

If we plan to cook a week schedule, we want to easily prepare a variety of dishes. Ease of the recipe preparation is related to the recipe similarity. The more the recipes have common ingredients, the easier the ingredients are prepared. The easy path to cook could be indicated by the shortest path between the recipes in the network. Figure 7 shows the shortest path from 'Chilled Kimchi Soup' to 'Grilled Deodeok' and it shows the example of various changes of dishes. 'Chilled Kimchi Soup' can be changed to 'Acorn Jelly and Rice' when 4 ingredients such as 'Rice', 'Egg', 'Carrot', and 'Welsh onion' are added in the recipe of 'Chilled Kimchi Soup'. Also 'Acorn Jelly and Rice' can be changed to 'Cucumber and Kimchi Rice' by adding 6 ingredients such as 'Beef', 'Soy Sauce', 'Sesame', 'Sugar', 'Garlic', and 'Pepper'.

<insert figure 7 here>

6. CONCLUSIONS

In this paper, we collected 897 recipes of Korea, China, Japan, Italia, and Thailand shared online, and explored the social relationship between recipes. To do this, we newly define the complexity of a recipe with the entropy measure based on information engineering. The complexity of a recipe can be classified as a difficulty of cooking preparation and a difficulty of cooking procedure. The entropy measure will be useful for evaluating the difficulty of cooking in objective criteria. For example, 'Garlic' is an ordinary ingredient, thus it is easy to prepare for cooking. But 'Licorice Root' is more difficult to prepare than 'Garlic', because 'Licorice Root' is an unusual

ingredient. We calculated the ingredient entropy based on the difficulty of ingredient preparation and then evaluated the complexity of the recipe by the summing of the ingredient entropies in the recipe. If any recipe needs many unusual ingredients such as 'Licorice Root', the recipe complexity will increase.

We measured the recipe similarity by the common ingredient entropies of two recipes. If two recipes include common ingredients having highly entropies, then the similarity between two recipes will be higher. The ingredient with high entropy would be indicate that it has more information about food.

We constructed the social network using the similarity between recipes. With this recipe similarity network, we found which recipes are clustering, which recipes are similar to other recipes, and is there a difference between some recipes. The similarity of recipes is defined in terms of the shortest distance in the network, Using this recipe similarity network, we can minimize the effort in ingredient preparation and cooking procedure when we planed the cooking schedule to prepare for many dishes. This means the same as the shortest distance between recipes in the recipe similarity network.

In future work, we plan to extend the recipe network to model the process that creates new cuisine from an original cuisine. This is related to the expansion model about the prediction and growing of social network. When the extended model of recipe network is implemented, it will be more easily predicted the generation of new fusion cuisines and also can be applied in developing new dishes.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A3A2042747).

REFERENCES

[1] Papadimitriou A, Symeonidis P, Manolopoulos Y. Fast and accurate link prediction in social networking systems. Journal of Systems and Software. 2012; 85:2119-32.

[2] Bliss CA, Frank MR, Danforth CM, Dodds PS. An evolutionary algorithm approach to link prediction in dynamic social networks. Journal of Computational Science.2014; 5:750-64.

[3] Ahnert SE. Generalized power graph compression reveals dominant relationship patterns in complex networks. Scientific Reports. 2014; 4(4385):1-6.

[4] Wang L, Li Q, Li N, Dong G, Yang Y. Substructure similarity measurement in Chinese recipes. Proc. of World Wide Web. 2008; 979-88.

[5] Teng C, Lin Y, Adamic L. Recipe recommendation using ingredient networks. Proc. of 4th Annual ACM Web Science Conference. 2012; 298-307.

[6] Freyne J, Berkovsky S. Intelligent food planning: personalized recipe recommendation. Proc. of Intelligent User Interfaces. 2010; 221-4.

[7] Ueda M, Takahata M, Nakajima S. User's food preference extraction for personalized cooking recipe recommendation. Proc. of 2nd Workshop on Semantic Personalized Information Management. 2011; 98-105.

[8] Kuo F, Li C, Shan M, Lee S. Intelligent menu planning: recommending set of recipes by ingredients. Proc. of the ACM Multimedia Workshop on Multimedia for Cooking and Eating Activities. 2012; 1-6.

[9] Shidochi Y, Takahashi T, Ide I, Murase H. Finding replaceable materials in cooking recipe texts considering characteristic cooking actions. Proc. of 9th ACM Multimedia Workshop on Multimedia for Cooking and Eating Activities. 2009; 9-14.

[10] Geleijnse G, Nachtigall P, Kaam PV, Wijgergangs L. A personalized recipe advice system to promote healthful choices. Proc. of 16th Intelligent User Interfaces. 2011; 437-8.

[11] Ahn YY, Ahnert SE, Bagrow JP, Barabasi AL. Flavor network and the principles of food pairing. Nature Scientific Reports. 2011;1-21.

[12] Ueda M, Asanuma S, Miyawaki Y, Nakajima S. Recipe recommendation method by considering the user's preference and ingredient quantity of target recipe. Proc. of the International Multi-Conference of Engineers and Computer Scientists. 2014; 1:12-4.

[13] Kim SD, Lee YJ, Kang SH, Cho HG, Yoon SM. Constructing cookery network based on ingredient entropy measure. Indian Journal of Science and Tchnology. 2015; 8(23):1-9.

TABLE 1. DATA REMOVAL AND REFINEMENT FOR EXTRACTING THE INGREDIENT NAMES

Step	Work	Method	Type of work	Contents
1	Data removal	Pattern matching	Removing the word, not the ingredient name	• Words that mean the amount: 'little', 'one', or 'two spoon'. • Words that mean the shape: 'mashed', 'cut', or 'frozen' • Symbols: '+', '=', or '-'.
2	Data refinement	Manually	Refining the word, not matching the name	• Same ingredients but different grammatical words. • Same ingredients, but similar words: 'Anchovy Soup', 'Anchovy Broth', 'Anchovy Sauce'
			Refining the word attached an adjective	• Ingredients attached the adjectives: 'Mashed Garlic', 'Sliced Garlic', or 'Whole Bulb of Garlic' is refined as 'Garlic'.
			Classifying the similar names	• Similar ingredients are classified as an ingredient prototype: 'Rice Cooking Wine', 'Clear Strained Rice Wine', and 'Cooking Wine' are classified as the 'Rice Wine'.

TABLE 2. THE NUMBER OF INGREDIENT NAMES AND COOKING VERBS IN COUNTRY RECIPES

Country	# of recipes	# of ingredients		# of verbs	
		Before the refinement	*After the refinement*	*Before the refinement*	*After the refinement*
Korea	669	1,089	421	1,187	110
Italia	85				
Japan	73				
China	52				
Thailand	18				
Total	897	1,089	421	1,187	110

TABLE 3. THE FREQUENCY AND THE RANKING OF THE INGREDIENTS

Ranking of ingredients						Ranking of verbs					
Top	*Ingredient*	*Frequency*	*Bottom*	*Ingredient*	*Frequency*	*Top*	*Verb*	*Frequency*	*Bottom*	*Verb*	*Frequency*
1	Garlic	53.96%	1	Persimmon	0.11%	1	Boil	51.39%	1	Harden	0.11%
2	Salt	44.82%	2	Licorice Root	0.11%	2	Mix	46.38%	2	Twist	0.11%
3	Welsh Onion	42.03%	3	Thistle	0.11%	3	Mince	40.58%	3	Thread	0.11%
4	Soy Sauce	41.92%	4	Leopard Plant	0.11%	4	Cut	39.91%	4	Strike	0.11%
5	Sesame Oil	36.34%	5	Costaria Costata	0.11%	5	Fry	35.45%	5	Extend	0.11%
6	Sugar	34.67%	6	Stonecrop	0.11%	6	Surround	32.89%	6	Beat	0.11%
7	Pepper	32.66%	7	Layou	0.11%	7	Cook	27.42%	7	Roll	0.11%
8	Onion	29.65%	8	Lobster	0.11%	8	Extract	25.75%	8	Dry	0.11%
9	Sesame	28.09%	9	Rum	0.11%	9	Sprinkle	21.07%	9	Dye	0.11%
10	Red Pepper	20.85%	10	Saposhnikoviae Radix	0.11%	10	Heat	17.50%	10	Erupt	0.11%

TABLE 4. HIGHER-RANKED RECIPE ENTROPY BY INGREDIENTS

Rank	Nation	Recipe	# of ingredients	Recipe entropy
1	Korea	Braised Dongtae Seafood	27	76.89
2	Italia	Vegetarian Lasagna	23	76.54
3	Italia	Lasagna	21	67.34
4	Thailand	Thai Salad	19	63.30
5	Italia	Italia Bread Salad	17	61.49
6	Italia	Greece Pasta Salad	19	60.67
7	Thailand	Thai Chicken Burger	17	59.70
8	Italia	Ribollita	17	59.61
9	Korea	Bulgogi Burger	20	58.46
10	Italia	Trout Asparagus Risotto	18	56.62

TABLE 5. LOWER-RANKED RECIPE ENTROPY BY INGREDIENTS

Rank	Nation	Recipe	# of ingredients	Recipe entropy
1	Korea	Boiled Potato	3	4.83
2	Korea	Rice with Egg and Soy Sauce	3	5.26
3	Korea	Steamed Egg	5	6.01
4	Korea	Candied Sweet Potato	2	6.41
5	Korea	Stir-fried Potato	4	6.54

TABLE 6. HIGHER-RANKED RECIPE ENTROPY BY VERBS

Rank	Nation	Recipe	# of verbs	Recipe entropy
1	Korea	Jeolla- Province Mosi-Songpyeon	28	82.08
2	Korea	Grilled Gochujang Chicken	31	80.84
3	Korea	Rice Porridge	26	61.19
4	Japan	Asian Orange Chicken	27	55.22
5	Korea	Chicken Raw-Ginseng Salad	20	52.91

TABLE 7. LOWER-RANKED RECIPE ENTROPY BY VERBS

Rank	Nation	Recipe	# of verbs	Recipe entropy
1	Korea	Healthy Silk-Fowl Soup	1	0.67
2	Korea	Gochujang	1	0.77
3	Japan	Dried Bean Curd Salad	1	0.77
4	Italia	Panettone	1	0.77
5	Italia	Greece Pasta Salad	2	1.55

TABLE 8. HIGHER-RANKED RECIPE SIMILARITY BY COMMON INGREDIENT ENTROPIES

Rank	Recipe	Recipe	# of ingredients	Similarity
1	Candied Sweet Potato	Sweet Potato Stick	4	1
2	Seasoned Bean Sprouts	Cooked Bean Sprouts	6	1
3	Balloon Flower Salad	Balloon Flower and Cucumber Salad	10	1
4	Medicinal Rice	Medicinal Rice by Electric Cooker	10	1
5	Pea Rice Gruel	Bean Gruel	3	0.91
6	Balloon Flower Salad	Balloon Flower Vegetable Salad	9	0.84
7	Balloon Flower Vegetable Salad	Balloon Flower and Cucumber Salad	9	0.84
8	Korean Bracken	Soaked Korean Bracken	7	0.81
9	Spit-Roasted Pine-Mushroom	Skewered Meatball	10	0.78
10	Spicy Marinated Crab	Spicy Crab	12	0.74

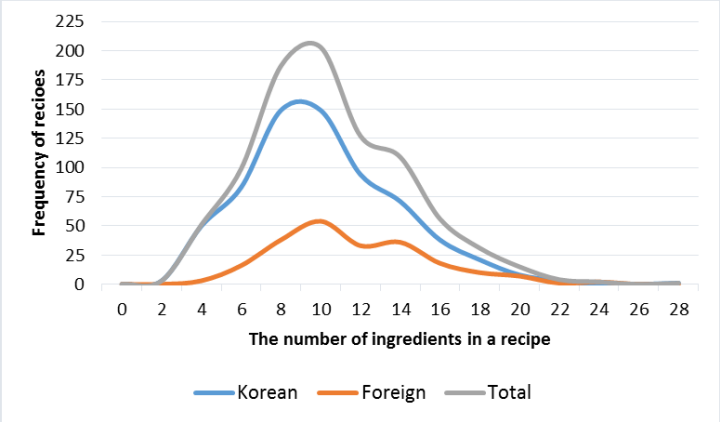

Figure 1. The number of recipes (y-axis) vs. the number of ingredients in a recipe (x-axis)

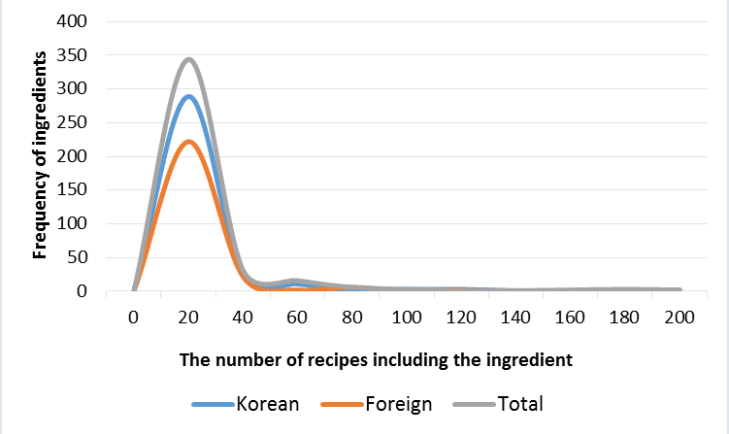

Figure 2. The number of ingredients (y-axis) vs. the number of recipes including the ingredient (x-axis)

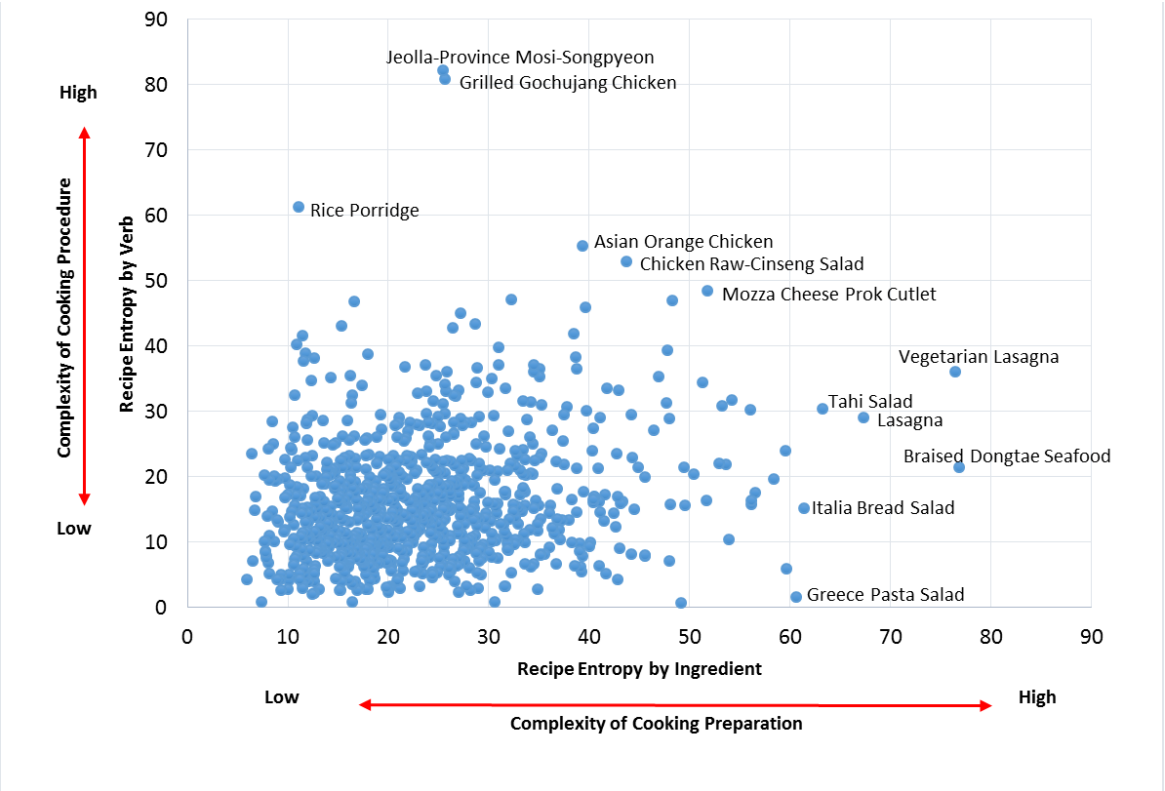

Figure 3. Distribution of recipe entropies by ingredient and cookung verb

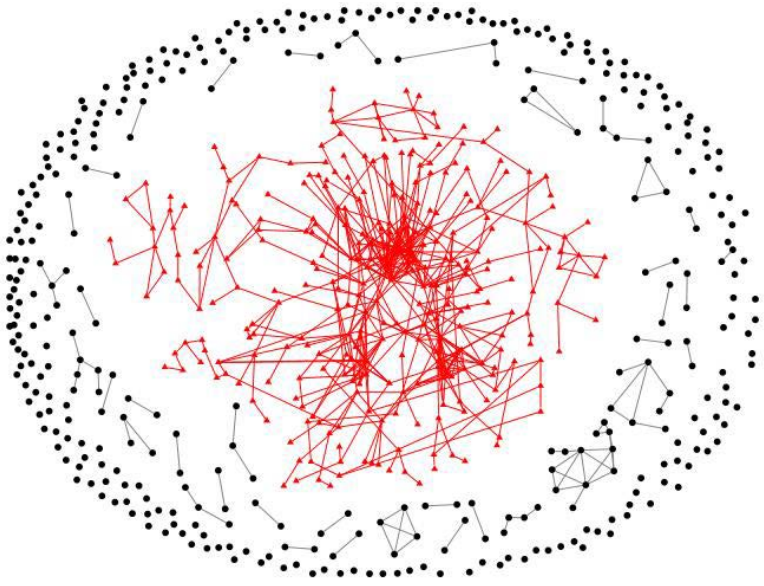

Figure 4. Korean recipe network when $\mathrm{Sim}(r_x, r_y) \geq t_0 (t_0 = 0.36)$.

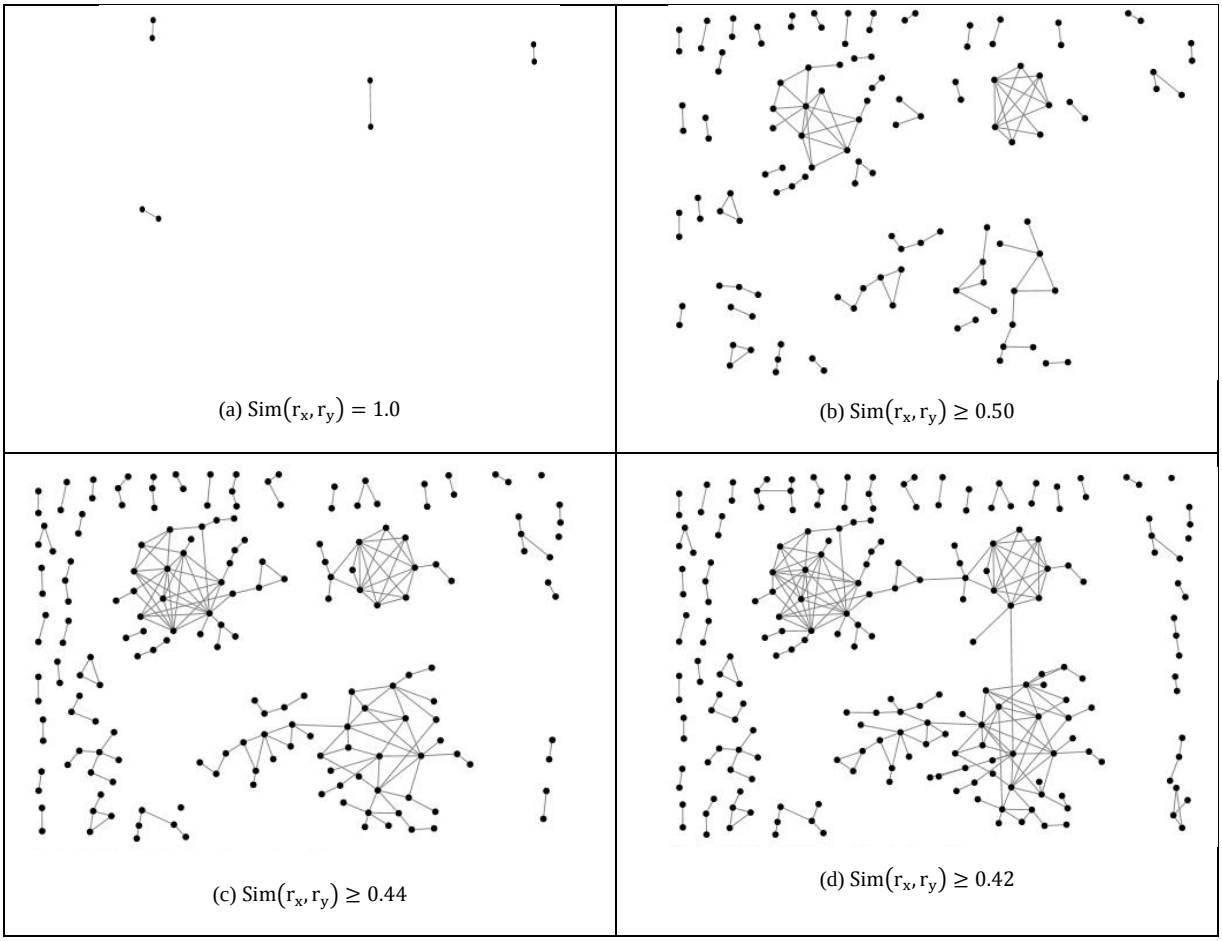

Figure 5. Connecting process of Korean recipe network when $\mathrm{Sim}(r_x, r_y) \geq t_0$.

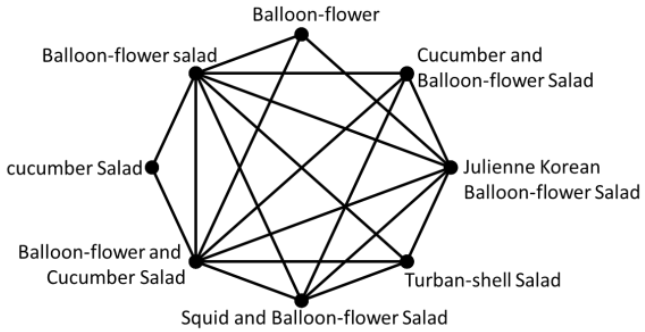

Figure 6. High related recipes in Korean recipe network when $\mathrm{Sim}(r_x, r_y) \geq 0.44$

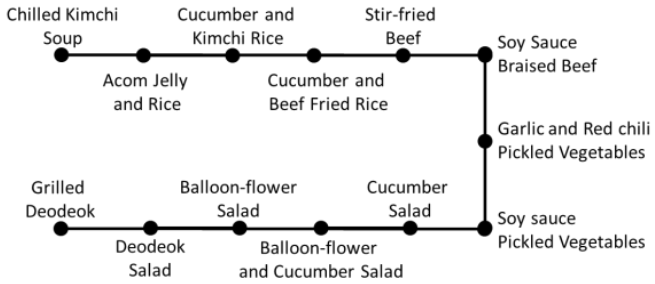

Figure 7. The Shortest path from 'Chilled Kimchi Soup' to 'Grilled Deodeok'.