

Flavor Based Recipe Recommendation

Team Foodies Progress Report

Haiyue Yin

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
hyin62@gatech.edu

Jidong Li

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
jli807@gatech.edu

Mo Shi

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
mshi61@gatech.edu

Shan Jing

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
sjing3@gatech.edu

Sixin He

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
she89@gatech.edu

Siyi Cao

Georgia Institute of Technology
North Avenue
Atlanta, Georgia 30332
scao34@gatech.edu

1 INTRODUCTION

From fire roasting, food preparation has become more than just making raw ingredients more edible. Nowadays people pay more attention to what they eat, and devote much more time in planning for delicious and healthy meals[9].

With the prevalence of online recipe sites such as *Cookpad*, *All-recipes* and *Yahoo! Recipe*, tremendous amount of recipe data is generated each day [18]. This prevalence leads to a growing demand for recipe recommendation systems.

Here we propose a novel recipe recommendation system that utilizes a combination of factors including user preferences, ingredient frequencies in recipes, and an ingredient flavor network. The goal of this project is to provide an easy-to-use interface that helps users decide on various potential tasty and healthy recipes.

2 SURVEY (LITERATURE REVIEW)

Recommendation Systems

Three strategies are mainly used by recommendation systems: content-based [15], collaborative and hybrid systems [14]. However, all of these strategies treat a recipe as an inseparable item [6]. Ueda et al. proposed a new recommendation system in which recipes are broken down into quantitative food ingredients, and user's food preferences are considered based on their browsing and cooking history. It was proven to be effective and reasonable to view a recipe as a repository of ingredient combinations and cooking methods [18][17].

There are some proposed problems of users' rating for recommendation system, such as the cold-start problem [4] and biased ratings[12]. Although Bayesian classifier may solve the cold-start problem [10], we still need to make good use of users' rating more carefully.

Recipe recommendation may consider a range of factors. Bianchini et al. consider prescription information as a key role for menu recommendation [3]. Both Karikome et al. and Ueta et al. try to make nutrition a key factor for their recommendation systems [8][19].

Food Pairing and Bridging

However, flavor is rarely addressed in the systems mentioned above. Food flavor is the main sensory property that influences food acceptance, and is usually the decisive factor for the choice of a particular product [7].

Various researches have been conducted on the patterns of chemical composition of ingredients in our daily recipes. The food pairing hypothesis states that people tend to put ingredients that share similar flavor compounds together. Ahn et al. also observe that most western cuisines follow the hypothesis by having ingredients that share similar flavors, while eastern cuisines avoid them [2][1]. Simas et al. further show a bridging phenomenon where an intermediate ingredient can smooth ingredients with contrast flavors. They observe that different regions display different combinations of levels of food pairing and food bridging [13].

3 PROGRESS OF PROPOSED METHODS

Data Collection

Using Spoonacular API we obtained about 50,000 recipes. Then we scraped the flavor information from Fenarolifis handbook of flavor ingredients.[13] [5]

Recipe Data Cleaning

The raw recipe data includes structured information such as cuisine type, image link, price, instructions, and a list of complete ingredients with ID's. The ingredient ID's were not directly available since Spoonacular uses a hash function for ingredient ID's. Fortunately Spoonacular's recipe data already contains pre-defined ingredients, and therefore we traversed all the recipes and obtained a list of ingredient ID's with the corresponding lists of ingredients names. Then we used the ingredient list to parse the recipes to obtain ingredient co-occurrences.

Flavor Data Cleaning

This was the most challenging part of data cleaning, since the Fenarolifis handbook is in PDF format. We converted it to txt using an online converter, and coarsely parsed the texts using NLP. With a combination of machine and manhours, we obtained a table of the chemical compounds and their corresponding natural occurrences

in important ingredient keywords that we named “essence”. We are yet to include the entries of botanic species in the book, which requires further manual processing.

Construction and Integration of Co-occurrence Network and Flavor Network

We matched the essence keywords to the Spoonacular’s ingredient names to obtain a mapping between chemical compounds and ingredients. The resulting flavor network is visualized in the visualization section.

Assume that we have n recipes and m ingredients which contains t chemical compounds in total.

Ingredient occurrence matrix K and flavor compounds matrix R were built to calculate the frequency of each two ingredients occurring in the same recipe and the number of shared chemical compounds between each two ingredient essence. K_{ij} means the j th ingredient is used in i th recipe or not, with corresponding values of 1 and 0. R_{ij} denotes whether the j th ingredient contains i th chemical compound. Then we have:

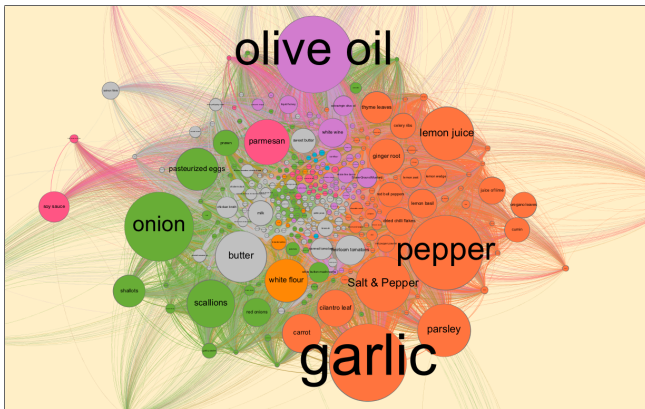
$$K^* = K^T K \Rightarrow K_{ij}^* = \sum_{p=1}^n K_{pi} K_{pj}$$

$$R^* = R^T R \Rightarrow R_{ij}^* = \sum_{p=1}^n R_{pi} R_{pj}$$

Through the above transformations, we have K_{ij}^* in K^* referring to the number of recipes that contain both the i th ingredient and the j th ingredient. R_{ij}^* refers to the number of shared compounds between the i th ingredient and the j th ingredient.

Database Construction and Data Visualization

Databases are built using SQLite3 for various data sets. One of the major table is for all the recipes, and the main data stored for it includes recipe names, ingredients, equipment and steps. Another important table is for storing data of flavor network. It consists by chemical compounds and the corresponding ingredients. Moreover, a co-occurrence table of ingredients is constructed. We used Gephi to visualize the ingredient network as shown below.



Above is a visualization of the flavor network. The node size corresponds to the prevalence of the ingredient, and the edge width

refers to the number of shared compounds. A simple modularity algorithm was performed on the graph, and one can observe that ingredient with similar flavors tend to be clustered together.

Data Analysis

After we got recipe data, we used Pointwise Mutual Information to get the co-occurrence rate between each two ingredients among our recipes, and use this PMI value to generate an ingredient network. PMI defined on pairs of ingredient (a,b):

$$PMI(a, b) = \log \frac{P(a, b)}{P(a)P(b)}$$

where:

$$P(a, b) = \frac{K_{ab}^*}{m}$$

$$P(a) = \frac{K_{aa}^*}{m}$$

$$P(b) = \frac{K_{bb}^*}{m}$$

The PMI gives the probability that two ingredients occur together against the probability that they occur separately. Complementary ingredients tend to occur together far more often than would be expected by chance.[16]

Recommendation

First, the users will be asked to enter a food ingredient they want to include in their desired recipe. After getting the input ingredient A from the user, a list of ingredients $A' = (A_1^T, A_2^T, \dots, A_n^T)^T$ will be found by using the index of PMI in the co-occurrence network sorted in descending order. Then the flavor network is used to filter the list A' by deleting the items who do not share the same chemical compounds with A and find out four top ingredients in A' for users to choose. The selected four ingredients should be those who have relatively high co-occurrences and also share some similar chemical compounds with A. Given the four ingredients, the users will be asked to select several ingredients that they like the most. These ingredients will be further used for recommendation.

Prototype interface

Foodies' paradise

sign in

Sign Up

Three Simple Steps:

☺ Enter one ingredient you choose to use

☺ Choose among these ingredients we recommend to make it just for you

☺ Click to generate and wait! We will do the magic for you

If you are ready, click

Next

Foodies' paradise

sign in Sign Up

Enter the ingredient you have in mind here

Foodies' paradise

sign in Sign Up

Ingredient you chose

Potato

Ingredients we recommend

Cheese

Tomato

Beef

Pepper

Next

Foodies' paradise

sign in Sign Up

Ingredient you chose

Potato Cheese Beef

Potato and Ground Beef Gratin

Ingredients

1 lb lean (at least 80%) ground beef

1 teaspoon salt

1/2 teaspoon pepper

3 tablespoons butter or margarine

1 small onion, chopped (1/4 cup)

3 tablespoons all-purpose flour

2 teaspoons chopped fresh thyme leaves

3 cups milk

3 cups shredded mild Cheddar cheese (12 oz)

6 medium white potatoes (2 1/2 lb), peeled, thinly sliced (6 cups)

Steps

1 Heat oven to 375°F. Spray 13x9-inch (3-quart) glass baking dish with cooking spray.

2 In 10-inch skillet, cook beef, salt and pepper over medium-high heat 5 to 7 minutes, stirring frequently, until beef is thoroughly cooked; drain.

3 Meanwhile, in 3-quart saucepan, melt butter over medium-low heat. Add onion; cook about 2 minutes, stirring occasionally, until tender. Stir in flour. Cook 1 to 2 minutes, stirring constantly, until bubbly; remove from heat. Stir in thyme.

4 Stir milk into onion mixture. Heat to boiling, stirring constantly. Boil and stir 1 minute; remove from heat. Stir in 2 cups of the cheese until melted.

5 Spread half of the potatoes in baking dish. Top with beef; spread remaining potatoes over beef. Pour cheese sauce over potatoes.

6 Cover with foil; bake 45 minutes. Sprinkle remaining 1 cup cheese over potatoes. Bake uncovered 15 to 20 minutes longer or until top is brown and bubbly and potatoes are tender. Let stand 5 minutes before serving.

4 PRELIMINARY RESULTS AND ANALYSIS

The preliminary results our group obtained are the co-occurrence network and flavor network, based on which we could achieve a reasonable recommendation of ingredients and recipes for the user. In particular, our group managed to recommend four ingredients given an input ingredient based on co-occurrence network and flavor analysis to the user, and therefore we can generate a recipe based on those ingredients.[11]

Below is an example of the result, with the first line being the queries, and the corresponding 4 recommendations.

Shrimp	Beef	Egg	Tomato
Clam	Carrots	Cottage	lettuce
Squid	Beef Broth	Bread Crumbs	Pita
Mussels	Onion	Fat Free Ricotta Cheese	Okra
Clam Juice	Veal	Ricotta Cheese	Seafood

Table 1: Time Line

The problem with our current algorithm is, the recommended ingredients may be too uncommon to give recipe recommendations. Also without tagging for recipes and ingredients, the recommendation may not satisfy usersfi need.

5 DESIGN OF UPCOMING EXPERIMENTS

Our group will adjust the number of ingredients obtained after the filtration using flavor table. Instead of using four final ingredients, five, six or even more ingredients could be used to be shown to the user to cover more ingredients.

Our group will experiment by adding tags to the recipes or ingredients to achieve a more purposeful recommendation. In particular, we can give recommendations of ingredients and recipes based on usersfi selected dish type and ingredient type.

We need to filter out uncommon ingredients when giving recommendation on ingredients. Therefore we have more flexibility when giving recipe recommendations.

6 LIST OF INNOVATIONS

Unlike popular recipe recommendation systems that either rely on ingredient frequencies or user ratings, we aim to integrate the two approaches and interpret the results with favor compounds analysis, which also allows us to generate customized recipes.

Since we are using flavor similarity between each two ingredients to give recommendation, we proposed a series of innovation recommendation methods that using flavor information. 1. We tried to use flavor similarity to substitute ingredients. 2. We tried to use flavor similarity as a threshold when recommending other ingredients to generate the recipe.

7 CONCLUSIONS AND DISCUSSION

We believe that with further modifications and ameliorations based on our proposed upcoming experiments, the recommendation system will give better results for the user and the user interface to be

built on the system will be easy and convenient to use.

8 PROJECT PLANNING

Original Planning and Task Assignmnets (Table 1 and Table 2):

Time	Task
Oct 16th	Gather and clean data
Oct 20th	Build data model
Oct 30th	Demo
Nov 6th	Build UI
Nov 10th	Progress report
Nov 13th	Add data analysis component
Nov 20th	Add user interactive function
Nov 27th	Improve features and functions
12/5	Final report

Table 2: Time Line

Task	Siyi	Mo	Sixin	Jidong	Haiyue	Shan
Proposal	✓	✓	✓	✓	✓	✓
Data Collection	○	○		○		○
Data Cleaning		○	○	○		
UI	○				○	
Database			○		○	○
Analysis	○	○		○	○	
Test		○	○			○
Finalization	○	○	○	○	○	○

Table 3: Task Assignments

Note: ✓ Completed ○Assigned

New Planning and Task Assignmnets (Table 3 and Table 4):

Time	Task
Nov 15th	Add data analysis component
Nov 20th	Build UI
Nov 23rd	Add user interactive function
Nov 27th	Improve features and functions
Dec 5th	Final report

Table 4: Time Line

Task	Siyi	Mo	Sixin	Jidong	Haiyue	Shan
Proposal	✓	✓	✓	✓	✓	✓
Data Collection	✓	✓		✓		✓
Data Cleaning		✓	✓	✓		
UI	○				○	
Database			✓		✓	✓
Analysis	✓	✓		✓	✓	
Test		○	○			○
Finalization	○	○	○	○	○	○

Table 5: Task Assignments

Note: ✓ Completed ○Assigned

Notes: All team members contributed similar amount of effort.

REFERENCES

- [1] Yong-Yeol Ahn and Sebastian Ahnert. 2013. The flavor network. *Leonardo* 46, 3 (2013), 272–273.
- [2] Yong-Yeol Ahn, Sebastian E Ahnert, James P Bagrow, and Albert-László Barabási. 2011. Flavor network and the principles of food pairing. *Scientific reports* 1 (2011).
- [3] Devis Bianchini, Valeria De Antonellis, Nicola De Franceschi, and Michele Melchiori. 2017. PREFer: A prescription-based food recommender system. *Computer Standards & Interfaces* 54 (2017), 64–75.
- [4] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems* 46 (2013), 109–132.
- [5] George A Burdock. 2016. *Fenaroli's handbook of flavor ingredients*. CRC press.
- [6] Jill Freyne and Shlomo Berkovsky. 2010. Intelligent food planning: personalized recipe recommendation. In *Proceedings of the 15th international conference on Intelligent user interfaces*. ACM, 321–324.
- [7] Henryk Jelen. 2011. *Food flavors: chemical, sensory and technological properties*. CRC Press.
- [8] Shihono Karikome and Atsushi Fujii. 2010. A system for supporting dietary habits: planning menus and visualizing nutritional intake balance. In *Proceedings of the 4th International Conference on Ubiquitous Information Management and Communication*. ACM, 56.
- [9] Marion Florence Lansing and Luther Halsey Gulick. 1920. *Food and Life*. Ginn.
- [10] Koji Miyahara and Michael J Pazzani. 2000. Collaborative filtering with the simple Bayesian classifier. In *Pacific Rim International conference on artificial intelligence*. Springer, 679–689.
- [11] Frank Restle. 1955. A theory of discrimination learning. *Psychological Review* 62, 1 (1955), 11.
- [12] Markus Rokicki, Eelco Herder, and Christoph Trattner. 2017. How Editorial, Temporal and Social Biases Affect Online Food Popularity and Appreciation.. In *ICWSM*. 192–200.
- [13] Tiago Simas, Michal Ficek, Albert Diaz-Guilera, Pere Obrador, and Pablo R Rodriguez. 2017. Food-bridging: a new network construction to unveil the principles of cooking. *arXiv preprint arXiv:1704.03330* (2017).
- [14] Janusz Sobecki, Emilia Babiak, and M S lanina. 2006. Application of hybrid recommendation in web-based cooking assistant. In *Knowledge-Based Intelligent Information and Engineering Systems*. Springer, 797–804.
- [15] Martin Svensson, Jarmo Laaksolahti, Kristina Höök, and Annika Waern. 2000. A recipe based on-line food store. In *Proceedings of the 5th international conference on Intelligent user interfaces*. ACM, 260–263.
- [16] Chun-Yuen Teng, Yu-Ru Lin, and Lada A Adamic. 2012. Recipe recommendation using ingredient networks. In *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, 298–307.
- [17] Mayumi Ueda, Syungo Asanuma, Yusuke Miyawaki, and Shinsuke Nakajima. 2014. Recipe Recommendation Method by Considering the User's Preference and Ingredient Quantity of Target Recipe. 2209 (03 2014), 519–523.
- [18] Mayumi Ueda, Mari Takahata, and Shinsuke Nakajima. 2011. Userfis food preference extraction for personalized cooking recipe recommendation. In *Workshop of ISWC*. 98–105.
- [19] Tsuguya Ueta, Masashi Iwakami, Takayuki Ito, and others. 2011. A Recipe Recommendation System Based on Automatic Nutrition Information Extraction.. In *KSEM*. Springer, 79–90.