Healthy Menus Recommendation: Optimizing the Use of the Pantry

Jefferson Caldeira  
Fed. Uni. of Campina Grande, Brazil  
jeffersonemanuel@copin.ufcg.edu.br

Leandro Marinho  
Fed. Uni. of Campina Grande, Brazil  
lbmarinho@computacao.ufcg.edu.br

Ricardo S. Oliveira  
Fed. Uni. of Campina Grande, Brazil  
ricardooliveira@copin.ufcg.edu.br

Christoph Trattner  
University of Bergen, Norway  
christoph.trattner@uib.no

ABSTRACT

We are often unable to plan menus ahead, thus making poor and unhealthy choices of meals. Besides healthy, one may want menus in which ingredients harmonize and cover well the available ingredients in the pantry. In this paper, we propose a novel multi-objective-based recommender of menus that features an optimal balance between nutritional aspects, harmony and coverage of available ingredients. We conduct experiments on real-world and synthetic datasets and show that our approach achieves the desired levels of nutrients, harmonization and coverage of ingredients.

CCS CONCEPTS

• Information systems → Recommender systems; • Applied computing → Consumer health; Health informatics;

KEYWORDS

Recommender systems; meal recommendations; nutrients optimization; multiobjective optimization

ACM Reference Format:

1 INTRODUCTION

There is a growing recognition that healthy food directly influences quality of life, by providing a sense of well being and happiness. However, eating healthily remains a challenge for many people. Among the possible reasons, the inability to conciliate the planning of healthy and tasty menus with the rush of everyday life is often present. By menu, we mean a set of meals, where each meal is comprised by a set of recipes and each recipe is comprised by a set of ingredients. In this paper, we will focus on lunch menus, leaving other types of menus for future work.

Besides fulfilling healthy nutritional standards, other properties should be considered before planning a menu, such as the harmony and easy availability of ingredients. In this paper, we introduce a novel lunch menu recommendation approach that considers all these properties simultaneously. Our algorithm receives as inputs the ingredients available at the user’s pantry and the number of portions desired, and recommends a lunch menu composed of a set of meals. The number of meals in the menu is decided automatically based on the availability of ingredients in the pantry. If a recommended menu has 7 meals, for example, the user could choose one different lunch meal for each day of the week. Each meal is composed of main dish, three side dishes divided into rice, beans, and pasta, salad, beverage, and dessert. An example of such a meal is roast chicken (main dish), garlic rice, black beans, spaghetti (side dishes), broccoli salad (salad), orange juice (beverage), and gelatin (dessert). We have chosen a meal setup particularly suited to the Brazilian food culture, where three of the authors reside, although it could be easily reconfigured to other food cultures.

We cast this as a multi-objective optimization problem, where standard nutritional indexes, harmonization and coverage of ingredients in the pantry are formulated as (possibly conflicting) objective functions. We use the Non-dominated Sorting Genetic Algorithm II (NSGA II) [7], which besides providing guarantees of convergence, feature diversity of recipes as an intrinsic property of the solution. We conduct experiments on real-world and synthetic data, and show that our approach is able to achieve an optimal balance between the desired level of nutrients, harmonization and coverage of ingredients.

1In this work we do not consider vegetarian meals, so a menu always contains some kind of meat.
Table 1: Notation used to describe our approach.

<table>
<thead>
<tr>
<th>Sym.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>set of nutritional components {proteins, carbohydrates, total fat}</td>
</tr>
<tr>
<td>( R )</td>
<td>set of recipes</td>
</tr>
<tr>
<td>( C )</td>
<td>set of recipes categories, e.g., main dish, side dish 1, side dish 2, side dish 3, salad, beverage and dessert</td>
</tr>
<tr>
<td>( M )</td>
<td>set of meals</td>
</tr>
<tr>
<td>( P )</td>
<td>set of pantries</td>
</tr>
<tr>
<td>( m \in M )</td>
<td>set of recipes where each recipe belongs to a different category of ( C )</td>
</tr>
<tr>
<td>( g_n(r) )</td>
<td>nutritional value of recipe ( r \in R ) regarding ( n \in N )</td>
</tr>
<tr>
<td>( f_n(m) )</td>
<td>nutritional value of meal ( m \in M ) regarding ( n \in N ), i.e., ( \sum_{r \in m} g_n(r) )</td>
</tr>
<tr>
<td>( I )</td>
<td>set of ingredients, e.g., rice and tomato</td>
</tr>
<tr>
<td>( I_{m} \subseteq R_{\geq 0} )</td>
<td>set of ingredients present in meal ( m \in M )</td>
</tr>
<tr>
<td>( I_{p} \subseteq R_{\geq 0} )</td>
<td>set of ingredients present in pantry ( p \in P )</td>
</tr>
<tr>
<td>( R_i \subseteq I_{m} )</td>
<td>set of recipes containing ingredient ( i \in I )</td>
</tr>
<tr>
<td>( V_\text{min}_n, \text{max}_n )</td>
<td>maximum and minimum values of the interval ( \text{P}_n ) resp.</td>
</tr>
<tr>
<td>( q_{m}(i) )</td>
<td>required quantity in grams of ingredient ( i \in I ) for preparing meal ( m \in M ), e.g., it is required 200g of rice for preparing a meal having garlic rice as side dish</td>
</tr>
<tr>
<td>( q_{p}(i) )</td>
<td>available quantity in grams of ingredient ( i \in I ) in pantry ( p \in P )</td>
</tr>
</tbody>
</table>

2 PROBLEM FORMALIZATION

For recommending menus featuring healthy nutritional values, good harmonization and coverage of ingredients in the user’s pantry, we first design specific objective functions for each one of these properties. Please refer to Table 1 for understanding the symbols used in this section.

Nutrition. A meal \( m \in M \) is composed of seven recipes, each one belonging to a different category of \( C \). For measuring if a meal complies to daily recommended lunch nutritional values, regarding the nutritional component \( n \in N \), we define \( \text{dist}_{n}(m) \in [0, 1] \) for computing the distance between the nutritional value of a meal and the range of nutritional reference values \( V_\text{ref} \). More formally, \( \text{dist}_{n}(m) \) is defined as:

\[
\text{dist}_{n}(m) = \frac{| f_n(m) - \text{min}_{V_n} | + | f_n(m) - \text{max}_{V_n} | - (\text{max}_{V_n} - \text{min}_{V_n})}{\text{max}_{V_n}}
\] (1)

If the nutritional values of the meal fall inside \( V_\text{ref} \), the function returns 0, otherwise it returns a value higher than 0. Values close to 1 mean that \( m \in M \) has a nutritional value close to the highest possible meal nutritional value in the dataset. In this work, we have used reference values provided by the Ministry of Health of Brazil (cf. Section 4). For example, suppose that a given meal \( m \in M \) has a protein value of 120g, i.e., \( f_{\text{protein}}(m) = 120 \), reference protein values in \( V_{\text{protein}} = [100, 150] \), and the maximum possible protein value of \( \text{max}_{V_{\text{protein}}} = 500 \). Applying \( \text{dist}_{n}(m) \) we have

\[
\frac{|120 - 100| + |120 - 150| - (150 - 100)}{500} = 0
\]

meaning that \( m \) complies with the protein reference values.

For later convenience, we will seek the meal that maximizes the inverse of the distance:

\[
\text{nut}_{n}(m) = 1 - \text{dist}_{n}(m)
\] (2)

Harmony. For measuring the harmony of a meal, we design an objective function where two ingredients are considered to harmonize well if they co-occur often in different recipes in the dataset. For each pair of ingredients in a meal, we compute the relative co-occurrence frequency of these ingredients considering all recipes in which they appear as ingredients. Eq. 3, defined in \( R_{\geq 0} \), formalizes this idea.

\[
\text{harm}(m) = \sum_{i,j \in I_{m}, i \neq j} \frac{|R_i \cap R_j|}{|R_i \cup R_j|}
\] (3)

where a value of 0 means that the ingredients of \( m \) do not harmonize at all.

Coverage. We seek to recommend menus that use, as much as possible, the available ingredients in the pantry. For that, we design a coverage function as the ratio between the available ingredients in the pantry and the required ingredients for composing the meal. Eq. 4, defined in \([0, 1]\), formalizes this idea.

\[
\text{cov}(m, t) = \frac{\min \left( \left( \sum_{i \in I_{m}, i \neq j} q_{p}(i) \right), \left( \sum_{i \in I_{m}} q_{m}(i) \right) \right)}{\sum_{i \in I_{m}} q_{m}(i) \times t}
\] (4)

where \( t \) is the number of portions required.

For example, suppose that a certain meal requires 200g of rice for one person, i.e., \( q_{m}(\text{rice}) = 200 \). If we consider 3 portions (\( t = 3 \)), we will require 600g of rice. If the pantry has available 600g of rice exactly, i.e., \( q_{p}(\text{rice}) = 600 \), then the coverage is maximum with a value of 1. If the meal requires less than what is available in the pantry, the coverage should also be maximum since we found all required ingredients in the pantry (that is why the \( \text{min} \) in the numerator of Eq. 4). If, however, the meal requires more than what is available, the function returns a value less than 1. So we seek meals that maximize this function.

Problem Statement. Given a set of ingredients \( I_{p} \) available at some pantry \( p \in P \) and the number \( t \) of portions required, our goal is to find the lunch meals that maximize all the aforementioned functions simultaneously, i.e.,

\[
\arg \max_{m} \left( \sum_{n \in N} \text{nut}_{n}(m), \text{harm}(m), \text{cov}(m, t) \right)
\] (5)

where \( k \) is the number of meals returned, calculated automatically by our algorithm (cf. Section 5). Notice that every time a meal that maximizes Eq. 5 is selected, the pantry needs to be updated accordingly and the coverage of the subsequent meals have to take these updated values into consideration.

3 RELATED WORK

Several related works have been proposed with the aim of recommending food to people. Trattner and Elsweiler [22] provides a good overview in this direction showing advances.
in recommender technology in the context of recipes, groceries or meals.

One of the earliest examples in this area are the works of Hammond [15] and Hinrichs [17] where case-based reasoning methods to recommend meal plans and recipes to people are introduced. More advanced algorithms employing content-based filtering and early stages of collaborative filtering include the works of Lawrence et al. [18], proposing a method to recommend groceries, and the work of Aberg [1], proposing a recommender method to recommend healthy recipes. Subsequently, more advanced methods and algorithms emerged for tackling different problems and aspects related to food recommendation. A good example in this direction is the work of Teng et al. [21] proposing a novel recipe recommender method based on ingredient networks. Other relevant studies include the work of Berkovsky and Freyne [4] proposing a method to recommend meals to groups of people; Harvey et al. [16] proposing a model that accounts for food selection biases; Ge et al. [13] proposing a method that leverages tags and latent factors to recommend recipes; Yang et al. [25] proposing the first constraint-based (with different types of diets) recommender, and the more recent work of Trattner et al. [24] proposing a novel method to recommend recipes to people in a cold-start scenario.

Other recent relevant works include Trattner and Elsweiler [23] showing the extent to which current recommendation algorithms are suitable for recommending healthy recipes. They were also the first to employ the WHO standards to recommend healthy recipes and meal plans.

Also of relevance are [12, 14] or [9] and [10] proposing a novel method to bring the ‘healthiness’ aspect into meal plans. In this direction, it is worth mentioning the works of Chifu et al. [5], proposing a solution using Particle Swarm Optimization to build healthy daily menu recommendations for elderly people; Agapito et al. [2] proposing a recommender system with focus on patients with chronic diseases; Cholissodin and Dewi [6] taking into account the family budget; and the work of Seljak [20] proposing a multi-objective approach for developing nutritionally and gastronomically adequate menus.

Summary & Contributions. The related works reveal several solutions available to tackle the food recommendation problem. Differently from us that recommend several meals grouped in a menu, most of these solutions focus on recommending recipes. Part of these solutions are concerned in recommending healthy food to people. Interestingly, we are not aware of any work that takes into consideration the ingredients that the user has available to prepare her food, as we do. Moreover, our approach takes into consideration a larger number of criteria in comparison to previous methods. In all, our contributions are summarized as follows:

- A novel method to recommend lunch menus considering, at the same time, reference nutritional values, harmony of ingredients, and coverage of pantry.
- Tailor designed objective functions for each property under consideration;
- The recommendation approach features easy to explain recommendations;
- Experiments showing that the recommended menu achieves the expected values of the desired properties.

4 NUTRITIONAL REFERENCE VALUES

The World Health Organization (WHO)\(^2\) is responsible for, among other things, setting norms and standards and assessing health trends world wide. It provides up-to-date references about healthy diets, which are used by many governments and institutions around the world for the definition of their own health policies.

The Ministry of Health of Brazil, for example, produced technical reports Ministério da Saúde [19] and a program called PAT - Programa de Alimentação do Trabalhador (Workers’ Nutrition Program) [8], containing nutritional reference values for healthy meals based on WHO, which we adopted in this work. In particular, we have adopted, considering an adult, a recommended energy intake of 2,000 kcal\(^3\), which results in a range between 600 and 800 kcal for the lunch meal. The reference values we have used, considering all nutritional components, are summarized in Table 2.

Table 2: Nutritional components reference values used in this paper.

<table>
<thead>
<tr>
<th>Component</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins</td>
<td>60g</td>
<td>120g</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>330g</td>
<td>600g</td>
</tr>
<tr>
<td>Total fat</td>
<td>90g</td>
<td>240g</td>
</tr>
</tbody>
</table>

5 MULTI-OBJECTIVE MENU RECOMMENDATION

Our approach receives as input:

- A set of recipes, categorized as: main dish (meat, chicken, pork, fish, etc.), three side dish categories (rice, beans and pasta), salad, beverage and dessert;
- A shopping basket, containing products normally used as ingredients for food preparation, as well as its quantities, representing the user’s pantry;
- A number of portions corresponding to the number of persons for whom the meals will be prepared.

\(^2\)http://www.who.int/

\(^3\)This is in line with the WHO reference values.
The decision of using a multi-objective approach is supported by the Pareto dominance concept, which is useful to compare different solutions across multiple objectives. For two candidate solutions, we say that one dominates the other if one solution is better than the other in at least one objective, when there is a tie in all others. For example, consider the objectives related to harmony and protein levels. Suppose that two meals are being compared, the first one with harmony and protein levels of 0.9 and 80g and the second one with 0.85 and 110g respectively. In this case, the first meal has a superior value of harmony, but there is a tie in protein since both meals fall into the healthy reference range for protein (60g to 120g). Thus, the first meal dominates the second with respect to these objectives. The set of all non-dominated solutions is called Pareto Front (or Pareto-optimal solutions) which represents the set of best possible solutions with respect to the objectives considered.

In our case, each candidate solution is a meal. In order to find the Pareto Front, it is necessary to compare each meal to every other meal in terms of the objective functions of interest, which leads to a combinatorial explosion problem. We employ the NSGA-II algorithm [7] for solving this problem. The reason for choosing NSGA-II is twofold: (i) it converges to Pareto-optimal solutions at lower complexity than a brute-force exhaustive search and (ii) it seeks to create diverse Pareto-optimal solutions. Concerning (i), the complexity of NSGA-II is $O(MN^2)$ where $M$ is the number of objective functions and $N$ the number of meals. Concerning (ii), this is particularly interesting because it may end up favoring meals that are diverse in terms of the recipes and ingredients used.

In the first iteration, NSGA-II randomly selects a parent generation of individuals (meals in this case). The size of this initial population is preserved across iterations. At each iteration, NSGA-II generates an offspring population through mutation (replacing one of the recipes) and crossing over (switching recipes between meals). The individuals are ordered by domination ranges in a process called Fast Non-dominated Sorting, in which each range contains individuals that do not dominate each other, but dominate the individuals in the next range.

If the number of individuals exceeds the population size, some individuals in the last domination range are selected in a process called Crowding Distance Sorting. This is done in order to spread the solutions along the Pareto Front, instead of concentrating solutions around similar objective values. This process is particularly useful in our context, since it can improve diversity of the meals along the iterations.

After a certain number of iterations, NSGA-II will yield a meal population that is Pareto-optimal. The final recommendation is formed by a set of meals extracted from this population, subject to the condition that a percentage of the ingredients needed are available in the pantry. In this work, we used 50% as this threshold, i.e., meals are recommended only if at least 50% of the required ingredients are available in the pantry. Notice that the selection of a meal causes a reduction of ingredients in the pantry, influencing the selection of the next meal. This selection is performed in a greedy manner, following the ranking provided by NSGA-II. Figure 1 summarizes the whole process of our recommendation approach.

### 6 Evaluation

In this section we present the evaluation of our approach. All code for the evaluation is publicly available online.

#### Data Collection and Preparation

The set of recipes used in this paper was collected from TudoGostoso®️, a Brazilian website of food recipes similar to Allrecipes.com®️. This is one of the most popular food websites in Brazil. A web crawler was implemented to collect the recipes, being executed from 07/19/2017 to 08/28/2017.

In total, 12,930 recipes were collected. Recipes about soups, alcoholic drinks, breads, snacks, etc., were discarded. We also collapsed the categories meat, chicken and fish, into main dish. In order to determine the nutritional information of a recipe, we first extracted the ingredients and their quantities present in the HTML. We then passed these ingredients to Tabela de Alimentos®️, a Brazilian website that receives an ingredient name (in Portuguese) and its quantity, and returns

![Figure 1: Pipeline of our recommendation approach.](https://github.com/JeffersonEmanuel/healthy-menus-recommendation)

---

6. [https://www.allrecipes.com/](https://www.allrecipes.com/)
nutritional information such as calories, carbohydrates, protein, total fat, fiber and sodium, about that ingredient. We wrote a script for automatizing this process. In this work, the relevant information are carbohydrates, protein and fat. For many reasons, such as typos and synonyms, ingredient names may not be found in *Tabela de Alimentos*. Thus, recipes containing any ingredient for which nutritional information were not found, were discarded. After this filtering, 741 recipes remained, divided into the seven categories used to compose the meal, as described in Table 3.

![Figure 2: Experimental results: our approach against meals formed by randomly chosen recipes, considering one portion.](image)

Table 3: Number of recipes in each category.

<table>
<thead>
<tr>
<th></th>
<th>Meat</th>
<th>Rice</th>
<th>Beans</th>
<th>Pasta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salad</td>
<td>255</td>
<td>69</td>
<td>48</td>
<td>84</td>
</tr>
<tr>
<td>Beverage</td>
<td>90</td>
<td>75</td>
<td>120</td>
<td>TOTAL</td>
</tr>
</tbody>
</table>

The dataset of shopping baskets, used to represent users’ pantries, contains 28 baskets. This dataset was collected from fellow graduate students and contain both shops made considering an entire family and people living alone. Each shopping basket is filtered in order to contain only food items (ingredients) and their respective quantities in the base unit of measure.

Personalization is achieved by recommending meals based on the ingredients available in user’s pantry. Due to the low number of pantries collected, a re-sampling was made in order to achieve the number of 1,000 shopping baskets. This was performed as follows. Two baskets are randomly sampled, where one of these baskets will receive a random number of new ingredients from the other basket (without repetition) where the quantities of each new ingredient are multiplied by a number in the range [0.5, 1.5]. So the quantities are shrunked, expanded or unchanged. We repeat this process until 1,000 baskets are produced.

Evaluation Protocol. The number of portions, representing the number of persons who will consume the meals, is another input necessary to the experiments. We run the experiments with the number of portions varying from one to four. Each recipe evaluated has its ingredients’ quantities multiplied by the number of portions, in order to determine the availability of ingredients in the pantry.

As baseline, we have used a random approach that will keep selecting random meals while at least 50% of the required ingredients are available in the pantry. This will serve to confirm that that our solution is not by chance.

Results. The experimental results show that, for a group of 1,000 pantries, our approach can recommend lunch meals that fit the daily nutritional requirements in what concerns proteins, carbohydrates and total fat, at the same time that provides good harmony and a good use of ingredients in the pantry.

Figure 2 shows the results. We show the box plots for 1 portion, that is, the meals are intended for one person, but the results are similar considering 2 to 4 portions. The blue horizontal lines in the left hand side of Figure 2 represent the recommended range for each nutrient, and T.E.V. means Total Energy Value.

First, in comparison to the random baseline, Wilcoxon tests showed that the distributions are significantly different for every tested objective function, with 95% confidence. Second, regarding the nutritional components, most of the recommended meals fall inside the recommended range, while the random approach has an erratic behavior as expected. Finally, the meals recommended by our approach present better values of harmony and coverage than random, as expected.
7 CONCLUSIONS

In this paper, we proposed a new approach for recommending lunch menus, in which a menu is defined as a set of meals. We cast the problem as multi-objective optimization problem where healthy nutrients, harmonization and coverage of ingredients in the pantry are considered simultaneously. To the best of our knowledge this is the first approach to consider this multitude of meals properties. We have used the NSGA II algorithm, a state-of-art multi-objective optimization solver, that also features diversity of recipes and ingredients as an intrinsic property of the solution.

Although our setup was tuned to fit Brazilian food culture and health standards, it can easily accommodate any food culture and health reference values. We conducted experiments on real and synthetic data that confirms the soundness and quality of our approach. As future works, we intend to compare our approach with state-of-the-art food recommender systems with a related purpose, specifically the works of Ahn et al. [3] and Cholissodin and Dewi [6]. Other purpose is to test the method with other recipe collections such as Allrecipes.com in the US or Kochbar.de9, one of the largest recipe platforms in Europe. Also, we intend to employ different standards as set by the World Health Organization (WHO) and Food Standard Agency (FSA) in the UK. Finally, we plan to conduct user studies in order to investigate qualitative aspects of our approach.

ACKNOWLEDGMENTS

This work has been supported by the project ATMOSPHERE (atmosphere-eubrazil.eu), funded by the Brazilian Ministry of Science, Technology and Innovation (Project 51119 - MCTI/RNP 4th Coordinated Call) and by the European Commission under the Cooperation Programme, Horizon 2020 grant agreement no 777154.

REFERENCES


9https://www.kochbar.de/