

# ユーザーの嗜好順序を利用した料理推薦システムの試作

An Implementation of a Recipe Recommendation System with User's Preference Order

岩上 将史\*<sup>1</sup> 伊藤 孝行\*<sup>2\*3\*4</sup>  
Masashi Iwakami Takayuki Ito

\*<sup>1</sup>名古屋工業大学大学院 情報工学専攻

Computer Science and Engineering, Nagoya Institute of Technology

\*<sup>2</sup>マサチューセッツ工科大学 スローン経営大学院

Sloan School of Management, Massachusetts Institute of Technology

\*<sup>3</sup>名古屋工業大学大学院 産業戦略工学専攻

Techno-Business Administration, Nagoya Institute of Technology

\*<sup>4</sup>科学技術振興機構 (JST) さきがけ 研究員

Researcher, JST PREST

To recommend daily recipes to users, in general we can't make use of the past user's preference because it is changing every day and every meal. In this paper, we present a new method for customizing recommendation results by using user's preference order to ingredients of recipes. Our system can recommend recipes which reflect personal taste in a few interaction.

## 1. はじめに

近年, allrecipes.com や cookpad.com などのユーザー参加型の料理レシピサイトが流行している。ユーザー参加型の料理レシピサイトは, 自分で作った料理情報を web サイトにアップロードでき, 作った料理に対してコメントがもらえるなどの特徴がある。しかし, 各ユーザーが自由に料理を投稿できるため, 投稿料理数が 70 万件を超えるなど, ユーザーの嗜好に合った料理を検索することが難しい。例えば, cookpad の料理検索システムで「カレー」と入力した場合, 料理名にカレーを含む料理が 25,221 件ヒットするため, 好みの料理を見つけ出すためにはさらに絞り込みを行う必要がある。また, ユーザーの多くはその日の気分により嗜好が変化し, 過去のデータが有効に活用できないため, 個人の嗜好を反映した推薦を行うことが難しい。そのため, web サイトにアクセスしてから画面を閉じるまでの間に嗜好を反映した推薦を行うには, ユーザーインタラクションから嗜好を獲得する仕組みが必要となる。

ユーザーインタラクションにより検索結果をカスタマイズする手法として, [山本 07] の再ランキング手法が有効な手法である。[山本 07] では, 検索結果に関連するキーワードをユーザーが強調および削除することにより検索精度を改善する。実験結果では, 3 回から 4 回のインタラクションにより検索結果の再現率が 8 割を超えることが確認され, 有効性が確認されている。

本論文では, 推薦された料理の材料をユーザーが使いたい(使いたくない)順にランク付けした結果を元にして推薦結果を再ランキングして料理を推薦するシステムを提案する。推薦結果をカスタマイズする際にユーザーに表示する材料に関しては, ユーザーが閲覧した料理を構成する材料をユーザーに提示し, ユーザーとのインタラクションから使いそうな材料を類推する。本システムにより, 少ない回数のインタラクションでユーザーの嗜好を獲得し, 個人の嗜好を反映した推薦を行うことが可能となる。

本論文の構成を以下に示す。第 2 章では, 本論文に関連した研究を紹介し, 本研究の位置づけを示す。第 3 章では, 本システムを作成するにあたって用いた料理間の類似度を定義する。第 4 章では, 本システムの構成を示す。第 5 章では, 本システムの実際の使用例を示す。そして, 第 6 章ではまとめと今後の課題を述べる。

## 2. 関連研究

本章では, 本研究に関連する研究として, エンティティ間のつながりを可視化する研究と, 対話に基づくインターフェースに関する研究を紹介する。

### 2.1 エンティティ間のつながりの可視化に関する研究

ソーシャルネットワーキングサービスの流行により, 近年エンティティ間のつながりに関する研究や, エンティティ間のつながりを可視化する研究が行われている。本研究でも, 料理間のつながりを可視化してユーザーに表示するインターフェースを採用している。人のつながりを可視化する研究の例として, ネットワーク作る君 [石田 07] が挙げられる。[石田 07] では, エンティティ間の関係の強さを定義するために, 以下の Jaccard 計数による式 (1) を用いている。

$$hit(a,b)/(hit(a) + hit(b) - hit(a,b)) \quad (1)$$

$hit(a), hit(b)$  はそれぞれエンティティ  $a, b$  の検索エンジンのヒット件数である。 $hit(a, b)$  は, エンティティ  $a$  と  $b$  のペアに対する検索エンジンのヒット件数である。式 (1) により, エンティティ間の関係の強さを計算し, 関係の強さがしきい値以上のエンティティ間の関係をネットワークとして表示している。また, [吉田 08] では「教えて goo!」の食べ歩きカテゴリの質問応答文を対象として, 関連語ネットワークの構築を試みている。つながりの可視化の応用例としては, あのひと検索 SPYSEE や mixiGraph などが挙げられる。

連絡先: 岩上将史, 名古屋工業大学大学院工学研究科情報工学専攻, 愛知県名古屋市昭和区御器所町, TEL:052-735-7968, FAX:052-735-7407, iwakami@itolab.mta.nitech.ac.jp

\*<sup>1</sup> ここでのエンティティは, 人の名前や物の名前をあらわしている。

## 2.2 対話に基づくインターフェースに関する研究

ユーザーの嗜好を考慮した料理レシピを推薦するにあたり、本システムでは、検索結果をユーザーが自分でカスタマイズする方法を採用している。ユーザーインタラクションにより検索結果をカスタマイズする手法として、再ランキング手法がある。[山本 07] では、ドラッグ&ドロップをユーザーが分類したい検索結果をシステムに伝える操作として導入し、検索結果に関連するキーワードをユーザーが強調および削除することにより検索精度を改善する。まず、ユーザーはクエリを入力し、検索結果を受け取る。そして、検索結果を閲覧していき、関係のなさそうな語を選択し、ごみ箱フォルダへとドラッグ&ドロップする。次に、関係のありそうな語を強調し、再ランキングされた検索結果を確認する。[山本 07] による実験結果では、3回から4回のインタラクションにより、検索結果の再現率が8割を超えることが確認され、ユーザーによる検索結果のカスタマイズの有効性が確認されている。

また、本研究では、ユーザーの嗜好順序を考慮して検索結果をカスタマイズしているが、優先順序を用いた検索の関連研究として、[塩澤 07] が挙げられる。[塩澤 07] では、キーワード検索に Dynamic Queries の手法を応用し、ユーザーがキーワードの優先度を対話的に変更でき、それに応じて検索結果の表示順が動的に変更されるインタフェースを提案している。具体的には材料ごとにスライダーを表示し、スライダーを調整することにより材料に優先度を付けている。

## 2.3 関連研究に対する本研究の位置づけ

本研究は、2.1 と 2.2 で紹介した手法を組み合わせている。まず、料理レシピサイトから収集した料理間の類似度を計算し、生成したネットワークをユーザーに提示する。そして、ユーザーが表示した料理を構成する材料をタグクラウド形式で表示する。ユーザーはタグクラウドから使いたい材料と使いたくない材料をランキング形式で配置し、システムは取得したランキング情報をもとに検索結果をカスタマイズする。本論文が既存研究と異なる点は、類似度の定義である。本システムでは、料理の関連ネットワークを生成する際、料理を構成する材料に加え、手順も加えて類似度を算出している点が異なる。また、[塩澤 07] との違いに関しては、[塩澤 07] はスライダーを用いて材料の優先度を決定しているが、本システムでは材料を使いたい(使いたくない)順に並び替えることによって材料の優先順序付けをしている。

## 3. 料理間の類似度

料理間の関連ネットワークを生成するためには、料理間の類似度を計算する必要がある。本論文での料理間の類似度は使用食材の類似度と調理手順の類似度を用いて計算している。

使用食材の類似度に関しては、本システムでは tanimoto 係数 [tanimoto 57] を用いる。tanimoto 係数は化合物の類似度を計算する際に用いられる指標であり、具体的な式は以下の式(2)で表される。

$$S_{ing} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (0 \leq S_{ing} \leq 1) \quad (2)$$

$A$  と  $B$  は料理  $A$  と料理  $B$  の各々の材料集合であり、記号  $|\cdot|$  は集合の要素数を表している。例えば、カレー {ルー, ジャがいも, にんじん} とカレーうどん {ルー, うどん} を比べる場合、カレーの材料数は3, カレーうどんの材料数は2, カレーとカレーうどんに共通する要素は {ルー} であるため、計算される類似度は、 $S_{ing} = 1/(3 + 2 - 1) = 0.25$  となる。

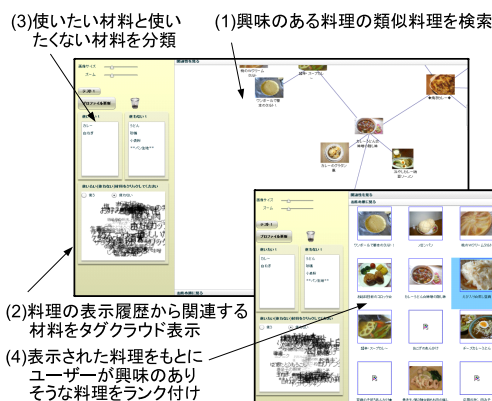


図 1: システムの構成

調理手順の類似度に関しては、編集距離を用いて計算する。編集距離とは、二つの文字列がどの程度異なっているかを示す数値である。具体的には、文字の挿入や削除、置換によって、一つの文字列を別の文字列に変形するのに必要な手順の最小回数として与えられる。“kitten”を“sitting”に変形する場合には、1.kitten 2.sitten (“k”を“s”に置換) 3.sittin (“e”を“i”に置換) 4.sitting (“g”を挿入して終了)のように最低でも3回の手順が必要とされるので、2単語間の編集距離は3となる。本システムでは、以上の編集距離を用いて、以下の手順で類似度を計算する。1. 自然言語で記述された調理手順を、文単位で区切る。2. 区切ったそれぞれの文を形態素解析し、名詞と動詞を取り出す。3. 取り出した名詞と動詞を用いて、文同士の編集距離を求める。4. 文同士の編集距離が閾値以下の場合には、二つの文章は同一であるとみなす。5. 二つの文章全体の編集距離を求める。具体例を示す。例えば、「じゃがいもと玉ねぎを炒める。ルーを入れる。」と「じゃがいもと玉ねぎを炒める。みりんを入れて煮込む。」という二つの調理手順が合った場合、一つ目の文章は{{じゃがいも, 玉ねぎ, 炒める}, {ルー, 入れる}}のように分解される。二つ目の文章は{{じゃがいも, 玉ねぎ, 炒める}, {みりん, 入れる, 煮込む}}のように分解される。文動詞の編集距離が2未満の場合を同一な文章とみなした場合、一つ目の文同士は同一となるが、二つ目の文同士は編集距離が2であるため、異なる文となる。従って、文章全体を比較した場合、二つ目の文章のみが異なるため、文章全体の編集距離は1となる。以上の編集距離を用いて、類似度を以下の式(3)のように定義する。

$$S_{proc} = \frac{1}{1 + E} \quad (0 \leq S_{proc} \leq 1) \quad (3)$$

二つの文章の編集距離  $E$  が0, つまり二つの文章が同一の場合は  $S_{proc}$  は1となり、二つの文章の編集距離  $E$  が大きい場合は  $S_{proc}$  は0に近い値をとる。

以上の式(2)と式(3)を用いて、類似度  $Sim$  を、

$$Sim = \frac{S_{ing} + S_{proc}}{2} \quad (0 \leq Sim \leq 1) \quad (4)$$

として定義する。

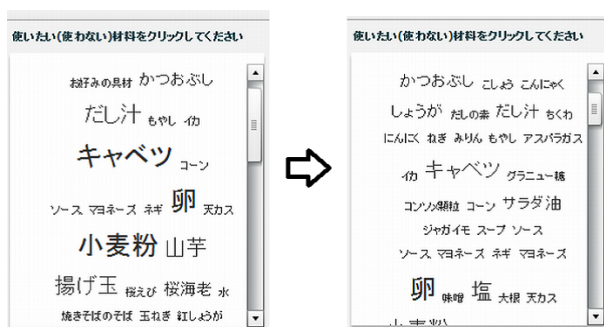


図 2: ユーザーの反応から材料のタグクラウドを動的に変更



図 3: 使いたい材料と使いたくない材料を分類

## 4. 料理推薦システム

### 4.1 システムの概要

本システムは、ユーザに料理名や使いたい材料を入力させ、ユーザの入力文から料理を検索して表示する。ユーザは気になる料理をクリックし、つながりのある料理を探索する。つながりを可視化することによりユーザに情報の整理を促すことが可能である。そして、獲得したユーザのクリック履歴や閲覧履歴からユーザが興味のある材料をタグクラウドで強調表示する。ユーザはタグクラウドから使いたい材料を順序付けし、得られた順序関係を利用して料理の検索結果をカスタマイズするシステムである。

本システムの流れを図 1 中の (1), (2), (3), および (4) の順に概説していく。(1) では、ユーザは興味のある料理をクリックし、tanimoto 係数と編集距離を用いて計算された類似料理を検索結果として得る。得られた結果はつながりのある料理としてネットワーク形式で表示される。(2) では、検索結果として得られた料理を構成する材料をタグクラウドで表示する。その際、ユーザの閲覧履歴からユーザが好みそうなタグを強調表示する。(3) では、ユーザはタグクラウドから使いたい材料と使いたくない材料を選択してリストに配置し、材料を使いたい(使いたくない)と思う順番にドラッグアンドドロップ操作により入れ替える。(4) では、順序付けされた材料を用いてシステムは料理の推薦順序を計算し、料理をランク付けしてユーザに提示する。

### 4.2 ユーザーインタラクションからのタグクラウドの調整

本システムでは、ユーザの閲覧行動を観測することによりユーザが興味がありそうな材料を予測する。嗜好を獲得する際、本研究では可視化されたネットワーク上でのユーザの料理の探索行為から嗜好が獲得できると考えた。具体的に観測する事象は、クリック行為、ブラウジング行為および選択した料理の閲覧時間である。上記の事象を考慮しユーザが興味を持った料理とそうでない料理を以下のように定義する。興味を持った料理: クリックして類似料理の検索に用いた料理、詳細情報を閲覧して長い時間見ていた料理。興味を持たなかった料理: 料理の探索を行っている最中に画面外に出てしまった料理。上記のユーザの反応を点数化してタグクラウドに表示される材料の大きさを調整する。具体的には、ユーザが興味を持った料理を構成する材料の点数を 1 増加し、興味を持たなかった料理を構成する材料の点数を 1 減少させる。そして、材料に付与された点数の大小に応じて表示する材料の文字フォントの大きさを調整する。詳細情報の閲覧時間に関しては、関

覧時間の分布が正規分布に従うと仮定して、平均値からの誤差が標準偏差の 3 倍以内の場合のみ、興味を持った料理とする。図 2 にタグクラウドが動的に変化する場合の実行例を示す。

### 4.3 嗜好に順序付けがある場合の tanimoto 係数

本システムでは、図 3 のようにユーザが使いたい材料を順序付けして検索を行うため、順序付けされた材料を用いて検索を行う場合に tanimoto 係数を拡張する必要がある。例えば、使いたい材料の順序が { ルー, じゃがいも, にんじん } だった場合、{ ルー, じゃがいも } と { じゃがいも, にんじん } では、通常の tanimoto 係数では値が同じになる。そこで、本システムでは、優先順序が高い材料に高い点数をつけるように変更する。具体的には、N 個の材料がある場合、点数は優先順位の高い順に  $2^0, 2^{-1}, 2^{-2}, \dots, 2^{-N+1}$  のようにつける。以上のように点数付けを行う理由は、上位に順序付けされている材料を最優先にするためである。1 位の材料が  $2^0$  点だとすると、 $2^0 - \sum_{k=2}^N 2^{1-k} > 0$  となり、1 位の材料は 1 位以外のすべての材料よりも優先される。

また、使いたい材料と使いたくない材料を考慮して料理をランク付けするため、

$$T = T_{good} - T_{bad} \quad (5)$$

として、 $T$  の値が高いものほどユーザが好む料理とする。ただし、 $T_{good}$  は使いたい材料集合との類似度で、 $T_{bad}$  は使いたくない材料集合との類似度である。具体例を示す。使いたいリストの順序の関係が「じゃがいも > にんじん > 玉ねぎ」で、使いたくないリストの順序関係が「ブロッコリー > アスパラガス」だった場合、ブロッコリーカレー { じゃがいも, にんじん, 玉ねぎ, ブロッコリー, カレールー } との類似度は、使いたい材料との類似度を計算する場合は  $T_{good} = (1 + 0.5 + 0.25) / ((1 + 0.5 + 0.25) + 5 - (1 + 0.5 + 0.25)) = 0.35$  となり、使いたくない材料との類似度は  $T_{bad} = 1 / (1 + 5 - 1) = 0.2$  となる。したがって、 $T = T_{good} - T_{bad} = 0.35 - 0.2 = 0.15$  となる。

## 5. 実行例と考察

試作したシステムによる実行例を示す。推薦に用いる料理データは料理情報サイトから取得した約 67 万件のデータである。本システムのインターフェースは料理間のつながりを表示する部分、材料のタグクラウド、ユーザの嗜好からのランクづけされた料理の推薦部に分けられる。入力部には、入力フォームを表示し、料理出力部には推薦料理を表示する。本実行例では、ユーザが「お好み焼き」と入力し、料理の関連ネッ





図 4: 使いたくない材料の順序が「ちくわ>にんにく>じゃがいも」の場合



図 5: 使いたくない材料の順序が「じゃがいも>ちくわ>にんにく」の場合

トワークを展開したとする。図 4 では、使いたくない材料の順序が「ちくわ>にんにく>じゃがいも」の場合の推薦である。そして、材料の順序を入れ替えて、じゃがいもを一番使いたくない材料として検索を行った結果が図 5 である。図 5 では、図 4 中の赤い枠で囲まれた「マッシュポテト入りお好み焼き」が推薦結果から除外されており、じゃがいもを使った料理に対して低いランク付けが行われていることがわかる。本システムでは、ユーザーのインタラクションに合わせて推薦される料理が変化するため、ユーザーの興味が変わっても対応可能である。例えば、カレーうどんからうどんに興味が変わり、さらに野菜をたくさん使った料理に興味が変わった場合も、ユーザーの使いたい材料と順序を考慮した推薦を行うことが可能である。また、嫌いな材料を除去した推薦が可能であるため、ユーザープロフィールを必要としない推薦がおこなえる。また、キーボード入力をほぼ必要とせず、クリックといったシンプルな入力により該当する料理を推薦可能である。

## 6. おわりに

本稿では、ユーザに料理同士の関係ネットワークを表示し探索を行わせ、ユーザーのインタラクションから嗜好を獲得することにより料理を推薦するシステムを試作した。また、材料を

用いた検索の際に tanimoto 係数と編集距離を利用し、材料同士の嗜好関係を考慮することでよりユーザーの嗜好を考慮した推薦を可能とした。しかし、材料同士の嗜好関係が増えてくると、検索条件が複雑になるため検索に時間がかかってしまう問題がある。また、検索条件を追加した方がユーザーの嗜好を獲得する際に役立つため、よりユーザーインタラクションが増やせるようにシステムを改良する必要もある。さらに、被験者による評価実験を行う必要もある。

## 参考文献

- [石田 07] 石田啓介, 松尾豊, 安田雪 ネットワーク作る君: Web マイニングを用いたネットワーク自動抽出システム, 人工知能学会全国大会論文集, 2007.
- [吉田 08] 吉田知訓, 間瀬心博, 北村泰彦 質問応答 Web サイトからの関連語ネットワークの自動抽出電子情報通信学会技術研究報告. AI, 人工知能と知識処理 108(119), 75-80, 2008.
- [山本 07] 山本岳洋, 中村聡史, 田中克己 Rerank-By-Example: 編集操作に基づく検索結果の網羅的閲覧日本データベース学会 Letters, Vol.6, No.2, pp.57-60, 2007 年 9 月.
- [tanimoto 57] T.T. Tanimoto. IBM Internal Report 17th Nov, 1957.
- [塩澤 07] 塩澤秀和, 三田村祐介: 食材の優先度を考慮した料理レシピの検索, 情報処理学会第 123 回 HCI 研究会, pp. 51-57, 2007 年 5 月.
- [中村 07] 中村聡史, 山本岳洋, 田中克己: Editable Web Browser: 編集操作の伝播によるウェブ閲覧支援, 情報処理学会第 123 回 HCI 研究会, pp. 59-66, 2007 年 5 月.
- [苺米 08] 苺米志帆乃, 藤井敦, 料理どうしの類似と組合せに基づく関連レシピ検索システム, 言語処理学会第 14 回年次大会発表論文集, pp.959-962, 2008.
- [多田 08] 多田和彦, 三石大, 佐々木淳, 船生豊. 媒介変数により個人の嗜好を反映可能なレシピ検索システムの構築と評価. 情報処理学会研究報告, No. 2002-DBS-126, pp. 137-144, 2002.
- [野間田 08] 野間田祐也, 星野准一: "GraphicalRecipes: レシピ探索支援のための視覚化システム", 芸術科学会論文集 Vol.7, No.2, pp. 43-54, 2008.
- [山本 09] 山本岳洋, 中村聡史, 田中克己 RerankEverything: ランキング結果を自由に閲覧するための再ランキングインタフェース, 第 17 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2009), 2009 年 12 月.
- [Bilenko 03] M. Bilenko, R. Mooney, W. Cohen, P. Ravikumar and S. Fienberg: "Adaptive name matching in information integration". Intelligent Systems, IEEE, 18, 5, pp. 16-23 (2003).
- [Muin 05] M. Muin, P.Fontelo, F. Liu, M. Ackerman: SLIM: An Alternative Web Interface for MEDLINE/PubMed Searches -A Preliminary Study, BMC Medical Informatics and Decision Making, Vol. 5, No. 1, 2005.