

# 主食材と主調味料を考慮した類似レシピクラスタリング

花井 俊介<sup>†</sup> 難波 英嗣<sup>††</sup> 灘本 明代<sup>†††</sup>

<sup>†</sup> 甲南大学大学院 自然科学研究科 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

<sup>††</sup> 広島市立大学大学院 情報科学研究科 〒 731-3194 広島市安佐南区大塚東 3-4-1

<sup>†††</sup> 甲南大学 知能情報学部 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

E-mail: <sup>†</sup>m1424007@center.konan-u.ac.jp, <sup>††</sup>nanba@hiroshima-cu.ac.jp, <sup>†††</sup>nadamoto@konan-u.ac.jp

あらまし 近年、ICTの発展に伴い、剽窃問題が顕著になってきており、大きな問題となりつつある。この剽窃問題はユーザ投稿型レシピサイトでも多く見受けられる。このような剽窃されたレシピが多数存在することは、倫理の問題のみならず、情報過多を招くことにより、ユーザの検索の妨げとなっている。そこで、我々は剽窃されたレシピの抽出のはじめの一步として、類似レシピのクラスタリング手法を提案する。我々のこれまでの研究より、類似レシピを判断する上で料理の主となる食材と味の決め手となる調味料の一致が重要であることがわかっている。そこで、本論文ではレシピデータからその料理の主食材、主調味料の推定を行い、推定した主食材、主調味料、そして食材の希少度に基づき、類似レシピをクラスタリングする手法を提案する。

キーワード 類似レシピ, クラスタリング, ユーザ投稿型レシピ

## 1. はじめに

近年、インターネットの発展に伴い、情報源の一つとしてWebを用いることが一般化している。しかし、Web上に存在する情報は容易に他者の文章をコピー&ペーストをすることができるため、剽窃が大きな問題となっている。剽窃問題はユーザ投稿型レシピサイトでも数多く見受けられる。例えば、テレビ番組で紹介されたレシピや書籍に掲載されているレシピを自分のオリジナルレシピとして投稿するケースや過去に投稿された他ユーザのレシピを少し改変して投稿するケースが見受けられる。料理レシピにおける著作権については著作権法関係裁判例<sup>(注1)</sup>によると、料理レシピなど物を作る際の手順を説明する表現物については、その作り方自体は「アイデア」であり「表現したもの」ではないため、著作物性の検討はあくまでもその説明等の表現物自体にあるとしている。「作品の著作権は、作品の作り方には及ばない」というのは著作権法の大原則とされているため、料理レシピに著作権法を適用することは困難であるとしている。また、料理の作り方は「キャベツともやしを炒める」や「水を入れ煮込む」のように表現方法が限られている。著作権法が適用されるようにするためには文章に詩的な表現のような独自の表現を用いることが求められる。しかし、料理レシピにおいて詩的表現などを用いることは実用的ではないため、限られた表現を用いることがほとんどである。このことはレシピが類似する要因の一つとなっている。類似したレシピが増加することにより、偶然または剽窃であるかの判断が困難となっている。また、剽窃されたレシピが多数存在することは、倫理の問題のみならず、情報過多を招くことにより、ユーザの検索

の妨げとなっている。

そこで、我々は剽窃されたレシピの抽出のはじめの一步として、類似レシピのクラスタリング手法を提案する。我々のこれまでの研究より、類似レシピを判断する上でレシピ同士の料理の主となる食材と味の決め手となる調味料の一致が重要であることがわかっている[1]。そこで、本研究ではレシピからその食材を欠かすとその料理として成り立たなくなる食材である料理の主となる食材を主食材、そのレシピ中で最も影響の強い調味料である味の決め手となる調味料を主調味料と呼び、その料理の主食材、主調味料の推定を行い、推定した主食材、主調味料、そして食材の希少度に基づき、類似レシピをクラスタリングする手法を提案する。

## 2. 関連研究

近年、剽窃の検出について数多くの研究が行われている。高橋ら[2]は、Web検索機能とn-gramと出現頻度を用いた剽窃評価機能、文章中の文字列が一定以上一致した部分にマークする剽窃箇所特定機能を用いて剽窃レポート発見支援システムを提案している。光原[3]はコピー&ペーストを制限した理想的なレポート作成モデルを提案し、このモデルに基づいてレポート作成支援システムを提案している。上田ら[4]は、最小操作回数による編集距離、および圧縮比率による情報距離の2つの手法を用いて類似性の判定し、剽窃レポートの検出を行っている。しかし、これらの研究はレポート等長文を対象としている。それに対し、レシピは比較的短文で書かれており、表現方法も限られている。そのため、既存の手法で抽出することは困難である。本研究では、剽窃レシピの抽出を行うため、まず、類似レシピの特徴の抽出を行い、ページ構造、単語の重要度に基づきクラスタリングすることで抽出を行う。

また、レシピ共有サービスの発展に伴い、レシピに関する研究が盛んに行われている。レシピ検索サイトにはCookpadや

(注1): 平成24年著作権法関係裁判例紹介, [http://www.jpaa.or.jp/activity/publication/patent/patentlibrary/patentlib/201306/jpaapatent201306\\_106135.pdf](http://www.jpaa.or.jp/activity/publication/patent/patentlibrary/patentlib/201306/jpaapatent201306_106135.pdf)

楽天レシピに代表されるユーザ投稿型レシピサイトやマークアップを用いた Google のレシピ検索システム<sup>(注2)</sup>、独自のレシピを掲載しているネスレバランスレシピ<sup>(注3)</sup>やぐるなびレシピ<sup>(注4)</sup>、ブログ形式で書かれたレシピブログ<sup>(注5)</sup>など、様々なレシピ検索サイトが存在する。しかしながら、いずれのレシピ検索サイトにおいても新着順や人気順などのソート機能やカテゴリで絞り込むといった検索方法が存在するが、類似したレシピを分類する機能は存在していない。

また、レシピ間の類似度に着目した研究も多数存在する。福本ら [5] は、食材の分量を食材の重要度、食品群ごとの平均分量によって重み付けし、それぞれのコサイン類似度を求め、コサイン類似度を合算した値をレシピ間類似度としている。苅米ら [6] は、材料、手順、に着目し、コサイン距離、DP マッチングを用いることでレシピ間の類似を算出している。また、Wang ら [7] は食材や調理手順を用いてレシピのグラフ化を行い、レシピ間に類似したサブグラフが存在することを示している。Li ら [8] はユーザの好みにあったレシピを抽出するため、料理レシピをグラフ化し、類似度を計算することでユーザごとの好みの食材、調理法を用いられている抽出する手法を示している。しかしながら、これらの研究では食材、手順の考慮はされているが、レシピの特徴を最も表していると考えられる、レシピタイトルについては考慮されていない。また、料理ごとの特徴的な材料についても考慮されていない。本研究では主食材や主調味料、ページ構造の各部ごとの役割や重要性、料理の種類ごとの食材の重要度を考慮し、レシピの分類を行う点が異なる。

食材に着目した研究も数多く存在する。Teng ら [9] は食材ネットワークを用いた食材推薦システムを提案した。食材間の関係を取得するために 2 つのタイプのネットワーク（食材補完、代替食材）の構築を行っている。Pinxteren ら [10] はレシピにおいて重要な特徴を明らかにし、抽出を行った。これらの特徴ベクトルに基づいて重み付き類似度を決定し、健康的なレシピに変更できるようにしている。志土地ら [11] はレシピテキストから食材名に対応する調理方法のマッチングを行うことによって代替可能な食材を発見する手法を提案している。Forbes ら [12] はレシピ推薦に行列因子分解法を適用した。実験結果から推薦精度の向上だけでなく、食材の代替や新たなレシピを作成することに有用であることを示している。しかし、これらの研究は食材の重要度に着目しているものの代替食材や代替レシピの推薦であり、類似レシピの抽出する我々の研究とは異なる。

調理手順の構造化に関する研究では、Kuo ら [13] はレシピサイト上に存在するレシピ間の共起関係を取得するためにレシピグラフを構築する手法を提案した。山肩ら [14] はレシピのフローグラフを作成することによって、複数のレシピから典型的な調理手順を抽出する方法を提案した。本研究はグラフモデルを用いず、ページ構造と単語の重要度に着目している点で異

なる。

### 3. 主食材、主調味料の抽出

我々のこれまでの研究 [1] より、類似レシピを判断する上でレシピ同士の主食材と主調味料の一致が重要であることがわかった。そこで、本研究では、レシピごとに主食材と主調味料の推定を行い、類似レシピを抽出する手法を提案する。

#### 3.1 材料リスト記入順の影響実験

ユーザ投稿型レシピサイトにおいて、人々は通常、メインとなる食材を各レシピの材料リストの最初に記入し、その食材がレシピの属するカテゴリに強く関連していることが分かっている [15]。しかし、この先行研究では材料リストに記入された材料すべてを用いて実験しているため、出現頻度では塩や胡椒といった一般的な調味料が上位となっている。そこで、本研究では主食材と主調味料を推定するために材料を食材と調味料に分け、材料リスト記入順の影響の実験を行う。実験手法を以下に示す。

(1) 「料理名」をクエリとし、レシピデータを取得する。

(2) 取得したレシピデータから、各レシピの材料リストを抽出する。

(3) 抽出した材料リストに出現する食材名と調味料名、出現順を我々の作成した食品データベースと照合することにより、抽出する。

(4) 食材名、調味料ごとに出現頻度、Tf-idf、材料リストの最初に記入された食材の出現頻度 (FI: First position of ingredient list) [15] を算出する。

データセット

実験に用いたクエリは一般的な料理名を用いた。実験で用いたクエリとレシピ数を以下に示す。

- オムライス 4659 件
- ハンバーグ 14274 件
- 肉じゃが 5184 件
- 白和え 2487 件
- 麻婆豆腐 3223 件

また、レシピデータにはクックパッドデータセット<sup>(注6)</sup>を用いた。実験結果

食材の実験結果を表 1、調味料の実験結果を表 2 に示す。出現頻度、Tf-idf、FI の上位 5 位を示している。表 1 より、食材の出現頻度ではその料理の一般的に用いられる食材が上位となっている。例えば、ハンバーグでは挽肉、玉葱、卵などが上位となっていることがわかる。しかし、挽肉はハンバーグの主食材であるが、玉葱や卵などは主食材として不適切である。また、ハンバーグのアレンジ料理である豆腐ハンバーグでは挽肉は用いられず、代わりに豆腐が用いられるため、常に挽肉が主食材であるとは限らない。Tf-idf では出現頻度と比べ、特徴的な食材が上位となっているが、ハンバーグの牛乳や片栗粉など主食材としては不適切な食材が多くなっている。一方、FI ではオム

(注2): レシピ検索- 検索サービス- Google <http://www.google.co.jp/landing/recipes/>

(注3): ネスレバランスレシピ <http://nestle.jp/recipe/>

(注4): ぐるなびレシピ <http://recipe.gnavi.co.jp/>

(注5): レシピブログ <http://www.recipe-blog.jp/>

(注6): クックパッドデータセット <http://www.nii.ac.jp/dsc/idr/cookpad/cookpad.html>

表 1 食材の出現頻度, Tf-idf, FI

クエリ: オムライス					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
卵	4934	牛乳	1680	米飯	2216
米飯	4463	鶏肉	1680	卵	889
玉葱	2648	人参	1576	鳥肉	455
鶏肉	1739	チーズ	1508	ネギ	227
牛乳	1738	玉葱	1444	玉葱	225
クエリ: ハンバーグ					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
挽肉	10962	牛乳	5237	挽肉	7976
玉葱	10495	豆腐	5125	豚肉	2187
卵	9253	片栗粉	4870	豆腐	1404
パン粉	7548	パン粉	4809	牛肉	1288
牛乳	4870	人参	4533	鳥肉	1215
クエリ: 肉じゃが					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
じゃがいも	4554	豚肉	1907	ジャガイモ	2058
玉葱	4138	こんにゃく	1879	豚肉	1097
人参	3382	牛肉	1865	牛肉	947
豚肉	1873	人参	1444	豚バラ	213
こんにゃく	1590	ささげ	1223	挽肉	195
クエリ: 白和え					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
豆腐	2412	人参	914	豆腐	1008
人参	958	ほうれん草	895	ほうれん草	278
ほうれん草	733	こんにゃく	712	春菊	80
こんにゃく	378	椎茸	503	飲料水	68
椎茸	199	ひじき	501	ニンジン	64
クエリ: 麻婆豆腐					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
豆腐	3388	青ネギ	1185	米飯	2216
片栗粉	2799	白ネギ	1179	卵	889
挽肉	2566	鶏肉	1157	鳥肉	455
にんにく	2316	生姜	880	ネギ	227
生姜	2134	玉葱	831	玉葱	225

表 2 調味料の出現頻度, Tf-idf, FI

クエリ: オムライス					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
ケチャップ	3604	バター	1651	ケチャップ	1485
食塩	3481	油	1594	食塩	768
コショウ	3010	ブイヨン	1539	バター	548
油	2248	サラダ油	1375	油	300
バター	2006	マヨネーズ	1345	ブイヨン	283
クエリ: ハンバーグ					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
食塩	11388	醤油	5247	食塩	6356
コショウ	10367	ケチャップ	5218	醤油	810
油	8104	ナツメグ	4910	ケチャップ	769
醤油	5052	酒	4870	ナツメグ	747
ケチャップ	4672	ウスターソース	4726	油	583
クエリ: 肉じゃが					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
油	4780	だし	1850	だし	982
醤油	4120	みりん	1696	醤油	825
砂糖	3618	酒	1646	砂糖	724
酒	2988	食塩	1631	油	577
みりん	2872	サラダ油	1595	酒	476
クエリ: 白和え					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
ごま	2190	味噌	911	ごま	827
砂糖	1605	食塩	893	味噌	330
醤油	1401	だし	862	砂糖	239
油	1347	油	826	醤油	202
食塩	1123	醤油	804	食塩	179
クエリ: 麻婆豆腐					
食材名	出現頻度	食材名	Tf-idf	食材名	FI
油	4754	鶏がらスープ	1182	味噌	1314
味噌	4391	砂糖	1149	ごま	356
醤油	2326	食塩	1122	油	276
ごま	1851	コショウ	1087	醤油	196
酒	1850	ごま油	1055	鶏がらスープ	179

ライスの米飯や卵, ハンバーグの挽肉や豆腐など主食材が多く取得されていることがわかる. FI の結果が最も良かった事より, 主食材の推定において食材の記入順を考慮することは有用であると考えられる. 次に, 調味料を対象とした結果を示す. 表 2 より, 出現頻度では一般的に調理に用いられる塩や胡椒といった調味料が上位となっており, これらは主調味料としては不適切である. また, Tf-idf では出現頻度と比べ, 特徴的な調味料が上位となっているが, オムライスではバター, 麻婆豆腐では鶏がらスープのように主食材としては不適切な食材が多くなっている. FI ではオムライスのケチャップや肉じゃがのだしや醤油, 白和えのごまや味噌のように主調味料が多く取得されている. このことから, 主調味料の推定においても調味料の記入順を考慮することは有用であると考えられる.

次に材料リストの 1 番目に出現する食材または調味料がその料理に対する主食材, 主調味料であるかを調査するため, 主食材, 主調味料が材料リストの何番目に出現するかを調査した. 調査にはクックパッドデータセットを用い, ランダムに 200 件

のレシピを取得し, 材料リストの何番目に主食材, 主調味料が出現したかを人手にて調査した. また, 食材, 調味料は分けて調査を行った. 記入順ごとの主食材, 主調味料の出現割合を表 3 に示す. 表 3 より, 主食材は材料リストの 1 番目に記入されることが多いことがわかる. しかし, 料理によっては主食材が複数存在する場合も考えられるため, 材料リストの 2 番目も出現することが多くなっている. 主調味料では主食材に比べ, 1 番目に記入された調味料が主調味料となる割合が少なくなっていることがわかる. 調味料は食材と異なり, 調理手順に記載する順に記入することが多く, 油や出汁, 塩といった調味料が 1 番目に記入されることが多くなったためであると考えられる.

#### 4. 類似レシピのクラスタリング手法

レシピの検索クエリとして, カレーや肉じゃがのような料理名をクエリとした検索と鶏肉やキャベツなどの食材名をクエリとした検索の 2 種類の検索が考えられる. マルハニチロホールディングスが行った「料理レシピに関する調査」[18]によると,

表 3 記入順ごとの主食材，主調味料の出現割合

主食材		主調味料	
記入順	出現割合 (%)	記入順	出現割合 (%)
1	85.0	1	44.5
2	29.5	2	21.5
3	9.5	3	10.0
4	5.0	4	2.5
5	1.5	5	2.0

どのような料理を作るか決定した状態でレシピ検索を行う人が全体の 26.4%であるのに対し，作る料理を決定していない状態でレシピ検索を行う人が全体の 56.5%となっており，作る料理を決定するためにレシピ検索を行っていると考えられる．そのため，レシピ検索を行う際のクエリは食材名が最も多く全体の 75.1%となっている．そこで，本研究では食材名をクエリとしたレシピ検索を対象とし，類似レシピをクラスタリングする手法を提案する．

#### 4.1 類似レシピのクラスタリング手法

我々のこれまでの研究 [1] [16] より，類似レシピの判断において，以下の 4 つの特徴があることが分かっている．

- 類似レシピの判断において画像は重要ではない．
- レシピタイトルに含まれる料理名，調理法名，食材名，調味料名の一致が最も重要である．
- 材料リストに含まれる主食材や味の決め手となる調味料の一致が重要である．
- シズルワード (“おいしさ”を連想させる食に関する言葉)[17] の一致は重要ではない．

これらの類似レシピの 4 つの特徴に基づき，類似レシピをクラスタリングする手法を提案する．提案手法の流れを以下に示す．

- (1) 「食材名」をクエリとし，レシピデータを取得する．
- (2) 取得したレシピデータから，各レシピのタイトル，材料リストを抽出する．
- (3) 抽出したタイトル，材料リストに出現する，料理名，調理法名，食材名，調味料名を我々の作成した食品データベースと照合することにより，抽出する．
- (4) 各レシピのタイトルから抽出した料理名，調理法名の単語を用いてクラスタリングを行う．
- (5) (4) のクラスタリングによって得た料理名，調理法名ごとの各クラスタに対し，料理ごとに主食材，主調味料の推定を行い，食材名，調味料名の出現する位置に基づいた特徴量を算出，ページ構造の重みを付与し，クラスタリングを行う．

#### 4.2 料理名，調理法名を用いたクラスタリング

類似レシピの特徴より，類似レシピの判断において，レシピタイトルの一致が最も重要であることが分かっている．また，レシピサイトは料理ごとにある程度分類されているが，実際に「食材名」をクエリとして検索を行った場合，検索結果には，様々な種類の料理レシピが混在していることがわかる．例えば「玉ねぎ ◯ ジャがいも」で検索した場合，その検索結果には「カレー」や「シチュー」といった様々な種類の料理が混在している．そこで，我々は類似レシピのクラスタリングのため

に，料理の種類ごとに検索結果を分類する必要があると考え，レシピタイトルに含まれる料理名，調理法名を用いてクラスタリングを行う．具体的には我々が作成した食材リストデータベースを用いて，レシピタイトルから料理名と調理法名を抽出する．抽出した料理名と調理法名の出現頻度を用いてクラスタリングを行う．この時クラスタリング手法には，我々の以前の実験結果 [19] より，Repeated Bisection [20] を用いる．

#### 4.3 料理名，調理法名，食材名，調味料名の出現位置を考慮したクラスタリング

最初のクラスタリングの結果，料理名，調理法名ごとのクラスタを取得した．しかしながら，同一種類の料理であっても類似した料理であるとは限らない．例えば，「玉ねぎ ◯ ジャがいも」をクエリとして，クラスタリングを行うと「カレー」のクラスタが生成される．しかし，このクラスタ内には「夏野菜カレー」「トマトカレー」「チキンカレー」「スパイシーチキンカレー」など様々なレシピが存在する．これらのレシピは玉ねぎとジャがいもを用いた「カレー」という同一種類の料理であるが，それぞれレシピの特徴が異なっており，類似レシピとは言えない．そこで，最初のクラスタリングで得たクラスタに対して，料理ごとに主食材，主調味料の推定を行い，推定した主食材，主調味料，食材の希少度，ページ構造の重みを付与し，クラスタリングを行う．

一方，我々の以前の実験より [16]，類似レシピの判断において，レシピタイトルに含まれる単語が最も重要であり，次にページ内の材料リストの領域内の単語が重要であり，調理手順の領域はあまり意識されていないことがわかった．そこで，2 回目のクラスタリングには，レシピタイトル，材料リストの中に含まれる料理名，調理法名，食材名，調味料名を用いる．

材料リストの記入順の影響調査より，主食材は材料リストの 1 番目に出現することが多いことが分かった．しかし，1 番目のみを用いると複数の主食材が存在する場合，対応することができない．また，主調味料では材料リストの 1 番目に出現する割合が少なくなっているため，出現順のみで主調味料を推定することは困難である．そこで，レシピタイトルに含まれる食材名，調味料名と材料リストの記載順，食材の使用分量を用いて主食材，主調味料の推定を行う．

また，料理ごとにその料理の特徴となる食材は異なると考えられる．例えば，「カレー」で一般的に使用されると考えられる玉ねぎやにんじんより，納豆やちくわのような一般的にはその料理に用いられないような食材の方が，その料理において特徴的であると考えられる．つまりは，ある料理に対してあまり使われない食材が使用されている場合はその料理の特徴を示すことがわかる．そこで本研究では，池尻らの提案する  $RF\_IIF$  (Recipe Frequency\_Inverted Ingredient Frequency) [21] の考えに基づいて，ある料理における食材の出現頻度と出現位置から食材の希少度を算出する  $S\_RF\_IIF$  を提案する． $S\_RF\_IIF$  の式を以下に示す．

$$S\_RF\_IIF_{i,m} = \alpha \log \frac{|R_m|}{|R_{i,t,m}|} + \beta \log \frac{|R_m|}{|R_{i,o,m}|} \quad (1)$$

$m$  は料理名を示し,  $i$  はある料理  $m$  に含まれる食材名または調味料名を示す. このとき,  $|R_m|$  はある料理  $m$  のレシピ数,  $|R_{i,t,m}|$  はある料理  $m$  のレシピ群中でレシピタイトルに食材  $i$  が出現するレシピ数,  $|R_{i,o,m}|$  はある料理  $m$  のレシピ群中で材料リストに食材  $i$  が出現するレシピ数である. また,  $\alpha$  は食材  $i$  がレシピタイトルに出現した場合の重み,  $\beta$  は食材  $i$  が材料リストに出現した場合の重みである. 我々はパラメータ決定の実験より  $\alpha = 1.0, \beta = 0.5$  とする. 最初のクラスタリングの結果のクラスタごとに主食材, 主調味料, レシピタイトルにおける料理名, 調理法名の出現頻度及び, 食材名, 調味料名の出現位置を考慮した希少度を用いて, 再度クラスタリングを行う. クラスタリング手法には最初のクラスタリングと同様 Repeated Bisection を用いる.

## 5. まとめと今後の課題

本論文では, ユーザ投稿型レシピサイトにおける類似レシピ抽出のために主食材, 主調味料の推定を行い, 推定した主食材, 主調味料, そして食材の希少度に基づき, 2段階のクラスタリングを行うことで類似レシピの抽出を行う手法を提案した. 1段階目のクラスタリングでは, レシピタイトルに含まれる料理名, 調理法名に基づきユーザの検索結果のクラスタリングを行い, 料理の分類を行う. 次に主食材, 主調味料, 食材の希少度, 単語の出現場所の重みを考慮し再びクラスタリングを行うことにより, 類似レシピの抽出を行う.

今後の課題として, 主食材, 主調味料の推定実験, また, 推定結果を考慮し, 類似レシピのクラスタリング実験を行う.

## 謝 辞

本研究の一部は JSPS 科費 26330347 及び, 私学助成金 (大学間連携研究補助金) の助成によるものである. また, 本研究を遂行するにあたり, クックパッド株式会社と国立情報学研究所が提供する「クックパッドデータ」を利用した. ここに記して謹んで感謝の意を表する.

## 文 献

- [1] 花井俊介, 灘本明代, 難波英嗣, “スパムレシピ抽出のための酷似レシピクラスタリング手法,” 情報処理学会研究報告, 2014-DBS-160(26), pp. 1-7, 2014.
- [2] 高橋勇, 宮川勝年, 小高知宏, 白井治彦, 黒岩丈介, 小倉久和, “Web サイトからの剽窃レポート発見支援システム,” 電子情報通信学会論文誌 D, 90(11), 2989-2999, 2007.
- [3] 光原弘幸, “Web を情報源とするレポート作成のためのコピー・アンド・ペースト制限とリフレクション支援,” 電子情報通信学会技術研究報告. ET, 教育工学 111.332: 1-6, 2011.
- [4] 上田和志, 富永浩之, “類似性に基づくレポート剽窃の検出ツールの改良とソースコードへの適用,” 電子情報通信学会技術研究報告. ET, 教育工学 110.453: 119-124, 2011.
- [5] 福本亜紀, 井上悦子, 中川優, “食材の重要度と食品群を考慮したレシピ間類似度の算出手法,” 第 4 回データ工学と情報マネジメントに関するフォーラム D9-2, 2012.
- [6] 苅米志帆乃, 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.
- [7] Wang, L., Li, Q., Li, N., Li, G. and Yang, Y. “Substructure similarity measurement in chinese recipes,” Proc. 17th

- International Conference on World Wide Web, pp. 979-988, 2008.
- [8] Li, Q., Chen, W. and Yu, L. “Community-based recipe recommendation and adaptation in peer-to-peer networks,” Proc. 4th International Conference on Ubiquitous Information Management and Communication, pp. 18:1-18:6, 2010.
- [9] Teng, C., Lin, Y. and Adamic, L. A. “Recipe recommendation using ingredient networks,” Proc. 4th International Conference on Web Science, 2011.
- [10] Pinxteren, Y. V., Geleijnse, G. and Kamsteeg, P. “Deriving a recipe similarity measure for recommending healthful meals,” Proc. 16th international conference on Intelligent user interfaces, pp. 105-114, 2011.
- [11] Shidochi, Y., Takahashi, T., Ide, I. and Murase, H. “Finding replaceable materials in cooking recipe texts considering characteristic cooking actions,” Proc. ACM multimedia 2009 workshop on Multimedia for cooking and eating activities, pp. 9-14, 2009.
- [12] Forbes, P. and Zhu, M. “Content-boosted matrix factorization for recommender systems: experiments with recipe recommendation,” Proc. 5th ACM conference on Recommender systems, pp. 261-264, 2011.
- [13] Kuo, F., Li, C., Shan, M. and Lee, S. “Intelligent menu planning: recommending set of recipes by ingredients,” Proc. ACM multimedia 2012 workshop on Multimedia for cooking and eating activities, pp. 1-6, 2012.
- [14] Yamakata, Y., Imahori, S., Sugiyama, Y., Mori, S. and Tanaka, K. “Feature extraction and summarization of recipes using flow graph,” Proc. 5th International Conference on Social Informatics, pp. 241-254, 2013.
- [15] Chung, Y. “Finding food entity relationships using user-generated data in recipe service,” Proc. 21st ACM international conference on Information and knowledge management, pp. 2611-2614, 2012.
- [16] 花井俊介, 灘本明代, “食材名をクエリとしたレシピ検索における酷似レシピクラスタリング,” 信学技報, vol. 114, no. 204, DE2014-31, pp. 47-52, 2014.
- [17] 大橋正房, 武藤彩加, 山本真人, 爲国正子, 汲田亜紀子, 洪澤文明, 小川裕子, “「おいしい」感覚と言葉食感の世代,” BMFT 出版部, 2010.
- [18] マルハニチロホールディングス, “~マルハニチロホールディングス、「料理レシピに関する調査」~, ” [http://www.maruhannichiro.co.jp/news\\_center/research/pdf/20130227\\_recipe\\_cyousa.pdf](http://www.maruhannichiro.co.jp/news_center/research/pdf/20130227_recipe_cyousa.pdf).
- [19] 花井俊介, 灘本明代, “酷似レシピ抽出のためのクラスタリング手法の提案,” DEIM Forum 2014 F8-6, 2014.
- [20] Zhao, Y. and Karypis, G. “Comparison of agglomerative and partitional document clustering algorithms,” Proc. SIAM Workshop on Clustering High-dimensional Data and its Applications, 2002.
- [21] 池尻恭介, 清 雄一, 中川博之, 田原康之, 大須賀昭彦, “希少性と一般性に基づいた意外性のある食材の抽出,” 日本ソフトウェア科学会「コンピュータソフトウェア」, vol.31, no.3, pp.70-78, 2013.