

食材利用履歴に基づき個人の嗜好を反映するレシピ推薦手法

Recipe Recommendation Method Based on Personal Use History of Foodstuff to Reflect Personal Preference

上田 真由美[✧] 石原 和幸[✧] 平野 靖[✧]
梶田 将司[✧] 間瀬 健二[✧]

Mayumi UEDA Kazuyuki ISHIHARA
Yasushi HIRANO Shoji KAJITA Kenji MASE

今日、料理レシピを推薦・紹介するサイトが多く存在するが、その多くは、登録日時やアクセス数の多い順にレシピを提供しており、利用者個人の嗜好を反映した推薦は行われていない。そこで本論文では、個人の調理履歴から算出した食材の得点と、レシピ内における食材の特異度に基づいてレシピの順位付けを行うレシピ推薦手法を提案する。本手法では、食材の得点付けを行う手法として、特徴的な単語の抽出に用いられるTF-IDF(Term Frequency - Inverted Document Frequency)の類推で、レシピを構成する食材の特徴の尺度となるFF-IRF(Foodstuff Frequency - Inverted Recipe Frequency)を提案する。本論文では、レシピへの得点付けの動作を確認するとともに、評価実験用システムを用いて適合率によって提案手法の有効性を示す。

There are many web sites recommending cooking recipe. They present cooking recipe in the order of entry date, access frequency or user's ratings, which can not reflect personal preferences. In this paper, we propose a novel frequency measure, Foodstuff Frequency - Inverted Recipe Frequency (FF-IRF). We apply the FF-IRF for personal foodstuff use history to reflect personal preference on to the recipe ranking. Furthermore, we evaluate the effectiveness of our recommendation method.

1. はじめに

夕食の献立について「今晚、何が食べたい?」と聞かれたことのある人や、聞いたことのある人は多いのではないだろうか。毎日料理をする人にとって、献立を考えることは労力を要する作業の一つであると考えられている。献立の決定に困った際には、家族に聞いてみたり、料理レシピの本やWebサイトを見たりして、献立を決めるきっかけとしている。日本テレビの「キューピー3分クッキング」[1]など長年放送されている有名な料理番組は存在するが、テレビ番組の場合、番組や企業の都合で紹介されるレシピが決定される。また、

個人がブログを使ってレシピを公開する場合も、公開者の都合でレシピは決定される。

代表的なレシピ検索サイトであるクックパッド[2]の2007年2月の利用者数は2,022,000人という結果がネットレイティングス株式会社によって報告されており[3], 料理レシピを提供するWebサイトの需要が高いことが伺える。クックパッドは、30万件以上のレシピが登録されており、食材をキーワードとして検索すると、新着順と人気順にレシピが提供される。例えば、このサイトで「豚肉」をキーワードに検索すると16,654件の結果が返される^①。登録ユーザが自由にレシピを登録することができるため、多くのレシピが登録されているが、類似のレシピが多かったり、少しでもその食材を使ったりしていれば検索結果となり、利用者にとっては「豚肉」料理とは思えない期待はずれの結果も多く返される。また、多くのサイトは登録順や名前順、カロリー順など表示順序を手動で切り替えることは可能であるが、すべての利用者に対して一律の結果を返すのみで、利用者個人の嗜好を反映させる手段はない。

そこで、我々は単純な食材をキーワードとして検索が行えるレシピ検索サイトに対して、利用者の調理履歴から抽出した個人の嗜好を反映させる手法を提案してきた[4]。本論文では、TF-IDF(Term Frequency - Inverted Document Frequency)における単語の特徴を尺度化する考えを食材の特徴の尺度化に応用したFF-IRF (Foodstuff Frequency - Inverted Recipe Frequency)を提案する。ここで、食材の利用頻度(FF)に個人の食材利用履歴を適用することで、個人の嗜好を反映させることが本手法の特徴である。本論文では、レシピへの得点付けの動作を確認するとともに、FF-IRFを用いた評価実験用料理レシピ推薦システムを構築し、提案手法の有効性を検証する。

2. 個人の嗜好を反映するレシピ推薦手法

2.1 提案手法の特徴

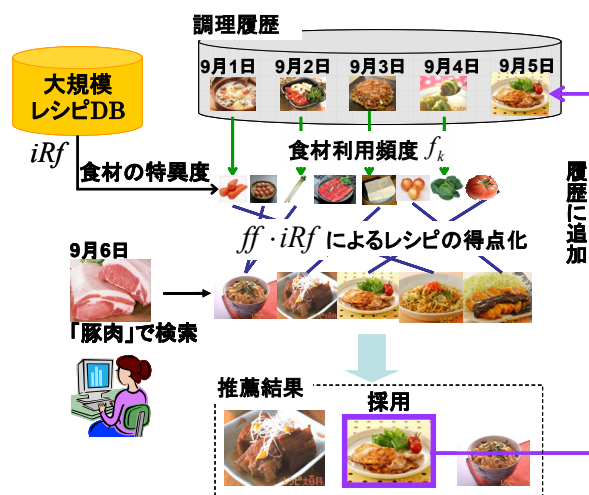


図1 食材利用履歴に基づくレシピ得点付けの流れ

Fig.1 Flow of the Scoring Method based on Use History of Foodstuff

[✧] 正会員 名古屋大学情報連携基盤センター
ueda@itc.nagoya-u.ac.jp

[✧] 非会員 名古屋大学大学院情報科学研究科博士前期課程
ishihara@arch.itc.nagoya-u.ac.jp

[✧] 非会員 名古屋大学情報連携基盤センター
{hirano, kajita, mase}@itc.nagoya-u.ac.jp

① 2008年1月28日現在

本手法では、個人の調理履歴から食材の得点を算出し、個人の嗜好として扱い、Web ベースのレシピ推薦システムが返す結果の順序付けを行う。以下に提案手法における得点付けの流れを示す(図 1)。

- (1) 利用者の調理履歴に含まれるレシピを食材単位に分解
- (2) 食材の個人利用履歴とレシピ群における得意度に基づいて各食材の得点 F_k を算出
- (3) 食材 k をキーワードとして候補となるレシピを検索
- (4) 候補となるレシピの得点 $Score_{i_m}$ を算出
- (5) 得点の高いものから順に推薦
- (6) 利用者が調理したものを調理履歴に追加
- (7) (1)に戻る

本手法は、調理履歴の中からシステムが利用者個人の嗜好を反映した「食材の得点」を算出するため、利用者は個人の嗜好をシステムに教示するために特別な操作を必要としない。また、推薦結果から調理したものを調理履歴にフィードバックするため、常に最新の嗜好を利用することが可能である。

2.2 食材スコアの算出

ユーザの調理履歴に含まれるレシピを食材単位に分解し、各食材の得点を算出したものを食材スコア F_k とする。食材スコア F_k はTF-IDFの考えを応用し、食材利用頻度(Foodstuff Frequency)と食材の特異度(Inverted Recipe Frequency)から算出する(式(1))。

$$\text{食材}k\text{の得点} = \text{食材利用頻度} \times \text{食材}k\text{の特異度} \\ F_k = f_k \times iRf_k \quad (1)$$

2.2.1 食材利用頻度

食材 k に対する食材利用頻度 f_k は、食材 k の出現頻度と調理経過日時に基づく経過重み $W_c = (c-1)/c$ から算出する(式(2))。

$$f_k = \begin{cases} \sum_{c=1}^t W_c \cdot 1 & (k \in i_c) \\ 0 & (k \notin i_c) \end{cases} \quad (2)$$

ただし、 $c(1 \leq c \leq t)$ は調理日からの経過日数を表し、 i_c は c 日前に調理したレシピを表す。ここで、頻繁に使用する食材は利用者が好きな食材であるとして頻度を合計しているが、過度に嗜好を反映しないように経過日数に応じた重み付けとした。また、重み W_c は、(1)毎日同じ食材となることは避けたい、(2)過去に用いた食材は利用者の嗜好を反映している、との考えから単調増加関数を用いた。例えば、過去 1 週間の献立において、3, 6, 7 日前に豚肉を用いた場合、豚肉の食材利用頻度は $f_{\text{豚肉}} = \sum_{c=1}^7 (c-1)/c \cdot 1 = 2.36$ となる。

2.2.2 食材 k の特異度

食材 k が料理を作る上でどの程度特異性があるかを示す特異度 iRf_k を以下の式で算出する(式(3))。

$$iRf_k = -\log_{10}(M_k/M) \quad (3)$$

ただし、 M はレシピデータベースに含まれる総レシピ数、 M_k はレシピデータベース内で食材 k を含むレシピ数を示す。ここで、レシピデータベースとしてクックパッドを対象とすると、 $M = 318,554$ 、 $M_{\text{豚肉}} = 16,654$ となり、 $iRf_{\text{豚肉}} = -\log_{10} 0.052279 = 1.28$ と計算できる。従って、豚肉の食材スコアは $F_{\text{豚肉}} = f_{\text{豚肉}} \times iRf_{\text{豚肉}} = 3.02$ となる。ある 1

週間の調理履歴をもとに、式(1)~(3)を用いて算出した食材の得点例を表 1 に示す。食材利用頻度と特異度の双方が高い「干し椎茸」の食材スコアは高くなり、食材利用頻度が低い「じゃがいも」の食材スコアは低くなっている。また、特異度 iRf_k が高い食材は、「生パスタ」($iRf_{\text{生パスタ}} = 3.52$)や「クジラ」($iRf_{\text{クジラ}} = 3.96$)などあまり使われていない食材となった。

表 1 食材の得点例
Table 1 Scores of Foodstuff

食材 k	食材利用頻度 f_k	食材 k の特異度 iRf_k	食材の得点 F_k
干し椎茸	3.11	2.37	7.36
人参	3.66	0.96	3.51
玉ネギ	2.16	0.80	1.72
ネギ	1.52	1.07	1.63
じゃがいも	0.5	1.23	0.62

2.3 レシピスコアの算出

ここで、上述した手順で得た食材の得点 F_k を加算することによって、推薦候補となるレシピ i_m の得点 $Score_{i_m}$ を算出する。ただし、レシピに含まれる食材の得点 F_k を単純に加算した場合、多くの食材を含むレシピの得点が高くなるため、食材の特異度 iRf_k で正規化を行う(式(4))。

$$Score_{i_m} = \frac{\sum_{k=1}^t F_k \cdot iRf_k}{\sum_{k=1}^t iRf_k} \quad (4)$$

通常、レシピを構成する食材数のレシピスコアへの影響を減少させるためには、レシピに含まれる食材の得点の単純和を食材数で割る手法が考えられる。しかし、食材によってレシピ内での重要度が異なることが考えられるため、すべての構成食材を同等に扱い食材数で割る手法ではなく、食材の特異度 iRf_k で正規化を行った[4]。例えば、「豚肉、じゃが芋、人参、そばつゆ、しらたき、玉ネギ」を材料とする肉じゃがのレシピスコアの計算は以下ようになる。

$$Score_{\text{肉じゃが}} = \frac{\sum F_k \cdot iRf_k}{\sum iRf_k} \quad (\text{ただし } k \in i_{\text{肉じゃが}}) \\ = \frac{F_{\text{豚肉}} \cdot iRf_{\text{豚肉}} + F_{\text{じゃが芋}} \cdot iRf_{\text{じゃが芋}} + \dots + F_{\text{玉ネギ}} \cdot iRf_{\text{玉ネギ}}}{iRf_{\text{豚肉}} + iRf_{\text{じゃが芋}} + \dots + iRf_{\text{玉ネギ}}}$$

この計算により、レシピに対して得点付けを行い、得点に基づいてレシピを順序付けることができる。

3. 提案手法の有効性検証

3.1 評価実験用システムの構築

提案手法の有効性を検証するため、評価実験用システムを構築した。料理レシピを推薦するシステムであるため、様々な利用者が想定される。そこで、利用者に複雑な操作を要求せずに単純な操作のみで自動的に個人の嗜好を抽出することが必要である。また、個人のコンピュータに対して複雑なインストール作業を要求することは不可能である。従って、本評価実験用システム利用時に利用者に要求する操作は、

Web ブラウザを介して

- (1) 料理したい食材を入力する,
- (2) 提供された料理レシピの中から気に入ったもの(調理するもの)を選択し調理する,

といった 2 種類の操作のみとする。(1)によってレシピデータベースに対する検索を開始し、システムが算出したレシписコアに従ってレシピを推薦する。また、(2)によって自動的に調理履歴として保存し、次回以降の食材スコア算出に利用する。この 2 種類の操作を行うのみで、利用者は意識することなく自動的に嗜好をシステムに伝えることができる。



図 2 評価実験用システムの画面例

Fig.2 Screenshot of the Experimental System

ここでは、クックパッドが提供する料理レシピを用いて推薦システムを構築した。1 章で述べたように、クックパッドに対して「豚肉」をキーワードにレシピ検索を行った場合、16,654 件の結果が返される。しかし、膨大な数のレシピが検索結果として返された場合、利用者がすべての検索結果を確認することは考えられない。そこで、本システムではクックパッドにおいて上位に提供される 90 件の結果を対象に得点付けを行い、利用者にはレシписコアに基づいて並び替えた上位 20 件の結果のみを提供する。図 2 に評価実験用システムの画面例を示す。左図は検索キーワード(食材)を入力し、推薦要求を行う画面である。利用者が推薦要求を行うと、推薦結果を提示する(図 2 右)。その中から調理したいレシピを選択すると、実際にレシピが表示される。

3.2 適合率による有効性の検証

3.1 節で述べた評価実験用システムを用い、家族分の食事を作る主婦 2 名と、自分用の食事を作る一人暮らしの男性 2 名の合計 4 名に利用してもらい、適合率により評価を行った。ただし、適合率は以下の式で計算する(式(5))。

$$\text{適合率} = \frac{\text{調理したいレシピの数}}{\text{推薦結果の数}} \quad (5)$$

被験者にあらかじめ 25 日分の調理履歴を記録してもらい、食材スコア・レシписコアの算出に利用した。各被験者には、
 (1) 個人の調理履歴を用いてレシписコアを算出し提示、
 (2) 一般的な調理履歴^②を用いてレシписコアを算出し提示、
 (3) 履歴は考慮せずクックパッドの返す順序で提示、
 の 3 種類の提示を行い、それぞれの提示に対して 5 つのキー

^② 一般的な調理履歴として、365 日分の料理レシピを提案している料理レシピ本[5]で提供されている 10 月 2 日～26 日のレシピを用いた。

ワードを用いて推薦要求を行った。被験者は、提示された料理レシピの中から、調理したいものを選択した。表 2～4 に「いわし」をキーワードとして推薦要求した際の、ある被験者に対して提示した 3 種類の提示例を示す。

表 2 個人の調理履歴を用いた場合
Table 2 Recipe Ranking (Personal Use History)

推薦順位	料理名
1	☆いわしと玉ねぎで蒲焼★
2	簡単☆鯛のムニエル★
3	焼きいわしとすっぱ野菜たち
4	イワシをのせて♪アボガドサラダ
5	ホワホワ☆お店の味☆
6	☆いわしの干物（一夜干し）の南蛮漬け☆
7	いわしの梅煮
8	いわしの酢煮、いわしの缶詰丼
9	いわしのマヨマスタード焼き
10	簡単☆いわしの蒲焼

表 3 一般的な調理履歴を用いた場合
Table 3 Recipe Ranking (General Use History)

推薦順位	料理名
1	簡単いわしのはんぺん！お好み焼き風
2	鯛の竜田揚げ ゆかり風味
3	生姜&大葉の香る、イワシのかば焼き
4	イワシの和風パン粉焼き♪
5	小骨も食べよう！いわしロールのトマト煮
6	簡単☆いわしの蒲焼
7	上品なお味に♪イワシの山椒煮
8	カリッ！の鯛ソテー☆簡単トマトソース
9	イワシのにんにく醤油煮
10	簡単☆鯛のムニエル★

表 4 調理履歴を用いない場合
Table 4 Recipe Ranking (Cookpad's Original Ranking)

推薦順位	料理名
1	超簡単！いわしのマリネ
2	イワシの和風パン粉焼き♪
3	イワシのつみれ鍋
4	鯛とキャベツのトマト煮込み
5	イワシのにんにく醤油煮
6	いわしのマヨマスタード焼き
7	簡単&本格的！トマトといわしのパスタ♪
8	鯛のさっぱり煮
9	いわしとごぼうのさっぱり煮
10	☆いわしの干物（一夜干し）の南蛮漬け☆

本実験により得た結果を図 3 に示す。被験者によって適合率の差は見られるが、どの被験者も個人の調理履歴を用いて順序付けを行った場合の適合率が高くなった。クックパッドは登録者が自由に料理レシピを投稿できるため、珍しい食材を使った料理も含まれている。各被験者の調理履歴を用いてレシピを順序付けることで、各被験者の調理履歴に含まれて

いない珍しい食材を使った料理のレシピスコアが低くなり、下位に順序付けられた結果、適合率が高くなったことが考えられる。文献[5]のレシピを調理履歴とした「一般的な調理レシピを用いた場合」も同様の理由から一般的な食材を用いた料理レシピが上位に順序付けられたため適合率が高くなったことが考えられる。図4に提示順序と選択回数の相関を示す。青が個人の調理履歴を用いて順序付けを行った場合、緑が一般的な履歴を用いた場合、ピンクが履歴を用いなかった場合を示している。調理履歴を用いなかった場合は、常に選択回数が少なくなっている。それに対し、被験者個人の履歴を用いた場合は、提示順序が上位のほうが比較的多く選択した結果を示している。

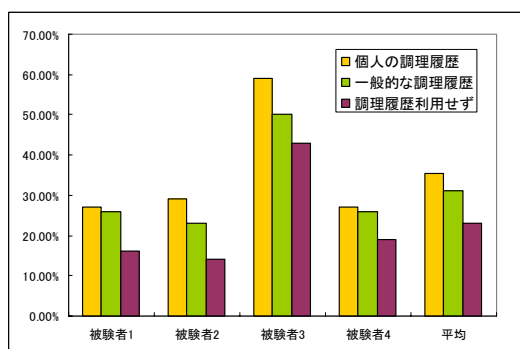


図3 適合率

Fig.3 Precision Factor

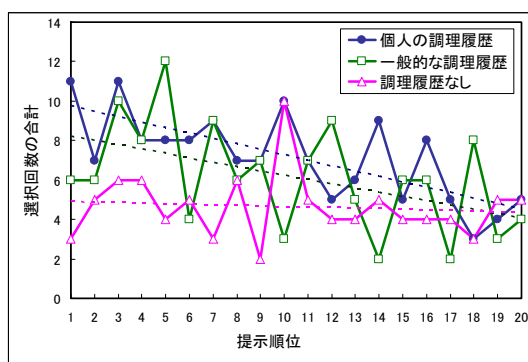


図4 提示順位と選択回数の相関

Fig.4 Correlation between Presentation Order and Number of Selection

4. まとめと今後の課題

本論文では、食材の利用履歴から食材に対して得点付けを行い、個人の嗜好を反映させた料理レシピの推薦手法について提案した。食材に対する得点付けを行う手法として、TF-IDFの考えを援用したFF-IRF(Foodstuff Frequency – Inverted Recipe Frequency)を提案した。これは、食材の得点の算出に、食材の利用頻度と調理を行う上での食材の重要度を用いるものである。特に、頻度による単純な特徴であるTF-IDFに対して、利用頻度を利用経過日数を導入し、嗜好の頻度を反映させるメカニズムを導入した点が新しい。

本論文では、評価実験用システムを構築し適合率による評価実験を行った。評価実験では各被験者に25日間にわたって晩御飯の調理履歴を記録し、(1)個人の調理履歴を用いてレシピの順序付けを行った場合、(2)一般的な調理履歴に基づい

てレシピの順序付けを行った場合、(3)調理履歴を用いなかった場合の3種類の提示に対して、調理したいか・興味がないかを判断してもらった。その結果、(1)の適合率が35.5%、(2)が31.3%、(3)が23.0%となり、提案手法の有効性を示した。

今後、被験者数やキーワードの種類を増やし実験を行う予定である。さらに、本論文では食材利用頻度 f_k の算出に用いた重み W_c として単調増加関数を用いたが、ガウス型関数となる W_c など食材利用頻度の利用について検討する必要がある。また、25日間の調理履歴を用いたため、すべての履歴を用いたが、履歴の有効期間についても検討する必要がある。

【文献】

- [1] “3分クッキング – お料理、レシピ満載！”，
<http://www.ntv.co.jp/3min/>.
- [2] “レシピ検索 No.1 / 料理レシピ載せるならクックパッド”，
<http://cookpad.com/>.
- [3] “Nielsen//Netratings Reporter Archives 2007年4月1日号”，
http://www.netratings.co.jp/hot_off/archives/NNR04012007.htm.
- [4] 上田真由美，石原和幸，平野靖，梶田将司，間瀬健二，
“食材利用履歴に基づき個人の嗜好を反映するレシピ推薦手法”，データベースをWeb情報システムに関するシンポジウム DBWeb2007, 3C-3, 2007.11.
- [5] “主婦の友生活シリーズ 新装版 365日の晩ごはん革命”，
主婦の友社，2001.

上田 真由美 Mayumi UEDA

名古屋大学情報連携基盤センター研究員。2003 関西大学大学院総合情報学研究科博士課程修了，博士(情報学)。同年大阪大学サイバーメディアセンター研究員，2005より現職。情報システムの個人化，学習支援システムの研究開発に従事。

石原 和幸 Kazuyuki ISHIMARA

名古屋大学大学院情報科学研究科博士前期課程在学中。2005名古屋工業大学システムマネジメント工学科卒業。情報推薦手法の研究開発に従事。

平野 靖 Yasushi HIRANO

名古屋大学情報連携基盤センター准教授。1999 名古屋大学大学院工学研究科情報工学専攻博士課程満了，博士(工学)。同年日本学術振興会特別研究員(PD)。2000 名古屋大学大学院情報工学専攻助手，2002 同大学情報連携基盤センター助手，2004 同助教授，2007より現職。

梶田 将司 Shoji KAJITA

名古屋大学情報連携基盤センター准教授。1995 名古屋大学大学院工学研究科情報工学専攻博士課程満了，博士(工学)。2002 名古屋大学情報連携基盤センター助教授，2007より現職。大学における教育・研究活動でのIT活用に関する研究開発に従事。

間瀬 健二 Kenji MASE

名古屋大学情報連携基盤センター教授。1981 名古屋大学大学院工学研究科情報工学専攻修士課程修了。同年NTT入社。1995~2002(株)国際電気通信基礎技術研究所研究室長。2002より現職。博士(工学)。コンピュータによるコミュニケーション支援の研究開発に従事。