

好き嫌いラベル付き食材分量を考慮したレシピスコア算出方式

中川明莉沙[†] 上田真由美^{††} 高畑 麻理[†] 中島 伸介[†]

[†] 京都産業大学 コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 京都大学大学院情報学研究科 〒606-8501 京都府京都市左京区吉田本町

E-mail: [†]g0946948@cc.kyoto-su.ac.jp, ^{††}mayumi@mm.media.kyoto-u.ac.jp,

^{†††}{g0846741,nakajima}@cse.kyoto-su.ac.jp

あらまし 近年、料理レシピの推薦手法に関する研究が数多く行われるようになった。我々は、ユーザの嗜好を考慮することが重要だと考え、先行研究として、食材に対する好き嫌いを考慮した料理レシピ推薦システムの提案を行ってきた。しかしながら、これまでの手法では、各レシピにおける食材の有無のみを考慮し、食材分量を考慮しないため、好きな食材が多くても少なくても、嫌いな食材が多くても少なくても、レシピスコアに影響を与えることはない。そこで、我々は、レシピ中の食材分量を考慮することで、利用者の感覚により近いスコアリングが可能となるレシピ推薦手法を提案する。本稿では、好き嫌いラベルが付与された食材分量のレシピスコアへの影響について考察すると共に、新たなレシピスコア算出手法について検討したので報告する。

キーワード レシピ推薦, 食材分量, 好き嫌い

Recipe Scoring Method by Considering Quantity of Ingredient Labeled Like or dislike

Arisa NAKAGAWA[†], Mayumi UEDA^{††}, Mari TAKAHATA[†], and Shinsuke NAKAJIMA[†]

[†] Faculty of Computer Science and Engineering, Kyoto Sangyo University.

Motoyama, Kamigamo, Kita-Ku, Kyoto-City 603-8555, Japan

^{††} Graduate School of Informatics, Kyoto University.

Yoshidahonnmachi, Sakyou-Ku, Kyoto-City 606-8501, Japan

E-mail: [†]g0946948@cc.kyoto-su.ac.jp, ^{††}mayumi@mm.media.kyoto-u.ac.jp,

^{†††}{g0846741,nakajima}@cse.kyoto-su.ac.jp

Abstract Nowadays, there are a lot of researches on the recommendation technique for cooking recipe. We have already proposed the cooking recipe recommender system based on considering users' preferences since we believe that it is important for developing such technique to consider user's preference. However, our previous work can consider only existence or non-existence of each ingredient in the cooking recipe but not the quantity of each ingredient. Therefore, the quantity of the ingredient has no influence on the recipe score for the recipe recommendation. Thus, we propose the recipe recommender system that can calculate the recipe score matching with user's preference by considering quantity of ingredient. In this paper, we discuss the influence of quantity of ingredient labeled like or dislike on recipe score and propose a new recipe scoring method.

Key words recipe recommendation, food quantity, like and dislike of ingredient

1. はじめに

生活習慣病の増加により、食生活改善が注目を集めている。しかし、毎日献立を考えることは労力を要する作業であることから、献立決定を支援するために栄養バランスや手持ちの食材を考慮したレシピ検索サイトの構築や、レシピ推薦手法に関す

る研究が多く行われている。既存の取り組みの多くは、食材や料理名を検索キーワードに用いて検索し、検索結果は登録日時やアクセス数、栄養バランス等によって順序づけて推薦を行っている。しかし、一般的に栄養素を摂取するためだけに食事をするのではなく、味や見た目を楽しまつつ食事をする事が多い。すなわち、栄養バランス等の条件を満たしていても、利用

者の好みに合わなければ調理されない。したがって、料理レシピを推薦する際には、利用者の好き嫌いを考慮することが重要である。

食事に関する好みは、人によって異なり、また、日々変化することが考えられるため、利用者自身があらかじめ嗜好を明確に記述しシステムに把握させることは、困難かつ煩雑であり現実的ではない。そこで、我々は利用者のレシピ閲覧・摂食履歴から自動的に食事に対する嗜好を推定し、レシピ推薦に用いる料理レシピ推薦手法の実現を目指している [1] [2]。特色としては、料理レシピを食材単位に分解し、レシピ閲覧・摂食履歴および食材の特異度を用いて個々の食材に対する嗜好を推定している。利用者の個々の食材に対する嗜好を得点化し、レシピに対して得点付けを行う。さらに、一般的に利用者は似た料理は連続して食べることは望まないと考え、食材、調理方法、料理名等を含め、類似した料理を連続して推薦しないように得点付ける。

しかしながら、上記のような手法では、料理中に含まれる食材の分量には着目しておらず、好き・嫌いと推定した食材が含まれるか否かでレシピに対する得点を算出していた。したがって、好きな食材が多くても少なくとも、嫌いな食材が多くても少なくとも、レシビスコアに影響を与えることはなく、同一の食材のみで構成されるレシピは同じ得点となっていた。同じ食材のみで構成される場合でも、好きな食材の分量が多い料理を好みやすく、嫌いな食材の分量が多い料理を避けやすいと考えられる。そこで、我々は、レシピ中の食材の分量を考慮することで、利用者の感覚により近いスコアリングが可能となるレシピ推薦手法を提案する。本稿では、好き嫌いラベルが付与された食材分量のレシビスコアへの影響について考察すると共に、新たなレシビスコア算出手法について検討したので報告する。

2. 関連研究

既存の料理レシピ検索システムの代表的なものとして、独自にコンテンツを作成しているレシピ大百科 [3] やマギーレシピサイト [4]、利用者による投稿レシピ用いるクックパッド [5] や Yahoo! レシピ [7]、楽天レシピ [6] など、さらに、既存のレシピ検索サイトをターゲットとしたメタ検索を行っている Google レシピ検索 [8] など様々なサイトがあげられる。

例えば、クックパッドは 100 万レシピが登録されており、1, 230 万人を超える月間利用者数があるように [9]、レシピ検索システムの需要の高さが伺われる。しかし、これらのレシピ検索システムは食材や料理名をキーワードに用いて検索を行い、登録日時やアクセス数の順に提供されており、個人の嗜好を考慮した順序付けは行っていない。

三野らはダイエット期間中のユーザのスケジュールに合わせたカロリーのレシピを推薦する手法を提案している [10]。日々の生活で起こりうるイベントに対して、カロリー消費・摂取の評価値を与え、期間内にダイエットを達成しうるカロリーのレシピを候補として抽出する。そのレシピ候補に対して塩分および野菜摂取量を制約条件とした線形計画を解くことにより、栄養バランスを考慮したレシピの推薦を実現している。また、苅

米らは栄養バランスがとれた献立の作成と食事記録の可視化を行っている [11]。この取り組みでは、「食品群別摂取量」に基づき、候補となる献立を食べた際の栄養素充足率を計算し、充足率に従って献立を順位付けている。さらに、栄養素の充足率および各食品群の充足率を可視化し利用者に提供することで、栄養バランスに関する問題発見を支援している [11]。Freyne らはコンテンツ型、協調型およびハイブリッド型の推薦手法を用いた料理レシピの推薦を行っている [12]。この取り組みでは、我々の手法と同様に、レシピを食材単位に分解して得点付けを行っている。しかし、レシピデータベースをあらかじめサーバ側に準備しておく必要があり、インターネット上のレシピデータの活用を考えている我々の手法とは異なる。さらに、彼らの取り組みでは、システムに嗜好や健康状態を伝達するため、各利用者があらかじめ大量のレシピに対して得点を付けておく必要がある。塩澤は既存のシステムで多く用いられている単純な食材名をキーワードとした検索では扱いきれない曖昧性に着目したレシピ検索手法を提案している [13]。この取り組みでは、Dynamic Queries 法をキーワード検索に応用し、単純な食材名の入力だけでなく、食材の優先度を動的に調整することを可能にしている。この取り組みは、利用者の曖昧な要求をシステムに伝えることを目指しており、利用者の好みを自動的に抽出しシステムに伝達することを目指している我々の取り組みとは異なる。

3. レシピ閲覧・摂食履歴を用いた嗜好の推定

3.1 献立決定の際に考慮すべき事項

偏った食生活は高血圧やメタボリックシンドロームなどの生活習慣病の一因となることが指摘されており、バランスの良い食生活が見直されている。しかし、栄養バランスを整えるために無理をして嫌いなものを食べる必要はなく、必要な栄養素を好きな料理から摂取できることが望ましい。そこで、献立決定時にどのような要素について考慮しているかを、20 代～40 代の男女 20 名に対してアンケートを実施したところ、以下のよう要素があげられた。

要素 1： 好き嫌い

要素 2： 栄養バランス、摂取カロリー

要素 3： 食材の有無、価格

要素 4： 調理の手間

要素 5： 気分

本研究では、上記の要素の中から、既存のシステムや研究で考慮されていない好き嫌いについて検討する。要素 5 の「気分」も既存のシステムでは考慮されていないが、主にレパートリーが豊富な上級者が気分によって献立を決定することができると考える。本システムは広く一般の利用者を対象としているため、まずは要素 1 の「好き嫌い」を対象とする。

3.2 食材に対する嗜好の推定

食事に対する嗜好の中で、本研究では食材に対する嗜好に着目する。食材に対する嗜好は、食材に対する「好き」と「嫌い」で構成され则认为、式 (1) で各食材の得点を算出し、嗜好として扱う。

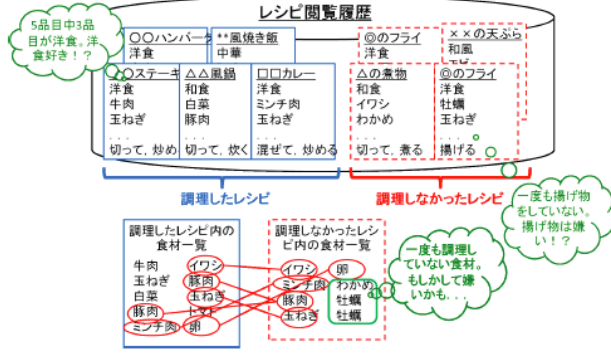


図 1 調理履歴を用いた好きな食材の推定

$$I_k = \alpha \cdot I_k^+ + (1 - \alpha) \cdot I_k^- \quad (1)$$

ここでは、食材 k の得点を I_k で表し、「好き」を I_k^+ 、「嫌い」を I_k^- とする。また、 α は「好き」と「嫌い」の嗜好に対する重みを表す。

3.2.1 好きな食材の推定

好きな食材の推定は、我々が従来から行っている利用者の調理履歴に含まれる食材の利用頻度と特異度を用いる手法 [1] に基づいて行う (図 1)。本手法では、「頻繁に使用する食材は利用者が好きな食材、または使用を避ける必要のない食材」と仮定する。

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 I_k^+ とする。食材スコアは TF-IDF の考えを応用し、食材利用頻度 (FF_k : Foodstuff Frequency) と食材の特異度 (IRF_k : Inverted Recipe Frequency) から算出する (式 (2))。

$$I_k^+ = FF_k \times IRF_k \quad (2)$$

このとき、食材 k に対する食材利用頻度 FF_k は、一定期間 D に調理された食材 k の頻度を用いて以下のように算出する (式 (3))。

$$FF_k = \frac{F_k}{D} \quad (3)$$

さらに、食材 k が料理を作る上で、どの程度特異性があるかを食材特異度 IRF_k として表し、式 (4) を用いて算出する。ここで、 M はレシピデータベースに含まれる全レシピ数、 M_k は対象食材 k を含むレシピ数である。

$$IRF_k = \log(M/M_k) \quad (4)$$

3.2.2 嫌いな食材の推定

食事に対する嗜好を考慮する際、好きな食材だけでなく、嫌いな食材についても考慮する必要がある。そこで、閲覧したが調理しなかったレシピに含まれる食材 k は「嫌いな食材、または使用を避ける食材である可能性がある」と仮定する。図 2 におけるエビがこれに該当する。

そこで、好きな食材の推定時と同様に、調理履歴およびレシピ閲覧履歴に含まれるレシピを食材単位に分解し、食材スコア I_k^- を式 (5) を用いて算出する。

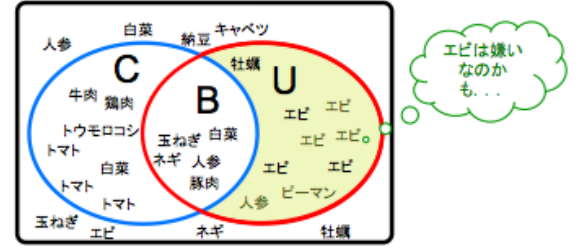


図 2 レシピ閲覧履歴と調理履歴を用いた嫌いな食材の推定

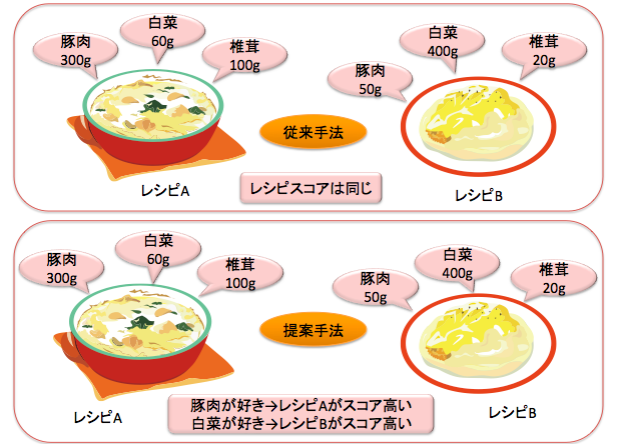


図 3 提案手法の基本的アイデア

$$I_k^-(x) = \begin{cases} 0 & \left(0 < \frac{2|U_k|}{|A_k|} \leq 0.5\right) \\ \left(\frac{2|U_k|}{|A_k|}\right)^x & \left(0.5 < \frac{2|U_k|}{|A_k|} \leq 1\right) \end{cases} \quad (5)$$

ここで、 $|U_k|$ は、図 2 中の集合 U に含まれる食材 k の個数、 $|A_k|$ はレシピデータベース中のすべてのレシピに含まれる食材 k の個数を表す。ただし、嫌いな食材であっても少量であったり、飾り付けのためなど、使ってしまうことも考えられる。そこで、どの程度の割合で避けた場合、その食材 k を嫌いであると判断するかを x で表す。

4. 食材分量を考慮したレシписコア算出方式

4.1 提案手法の基本的アイデア

従来手法と提案手法について比較をする。例えとして、レシピ A は豚肉 300g、白菜 60g、椎茸 100g 含み、レシピ B は豚肉 50g、白菜 400g、椎茸 20g 含むとする。

従来手法では、3.2 節で示したように、ある食材 k がレシピ中に含まれているか否か、に基づいてレシписコアを算出する。そのため、食材の分量にかかわらず好きなものが少量であってもレシписコアは高くなり、嫌いなものが少量であってもレシписコアが低くなってしまふ。具体例を図 3 上図に示す。豚肉が好きなあるユーザに対して、直感的には、豚肉を多く含むレシピ A を上位に提示すべきと考えられるが、従来手法ではレシピ A も B もレシписコアは同じになるため、同順位で提示する。同様に、椎茸が嫌いなあるユーザに対して、椎茸が少量であるレシピ B を上位に提示すべきと考えられるが、従来手法で

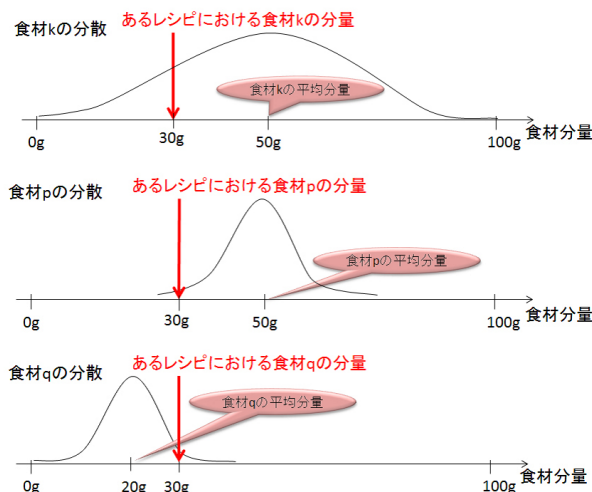


図 4 食材分量の分布と平均値の例

はレシビスコアは同じになり、同順位での提示となる。

これに対し、提案手法では、食材を含むか否かだけでなく、好き・嫌いと推定した食材 k の分量を考慮し、レシビスコアを算出する。すなわち、好きと推定した食材を多く含むレシピはレシビスコアを高く、嫌いと推定した食材を多く含むレシピのレシビスコアを低くする。具体例を図 3 下図に示す。すなわち、豚肉が好きなあるユーザに対して、従来手法ではレシピ A、B とともにレシビスコアは同じであったが、提案手法では豚肉 (好きな食材) を多く含むレシピ A のレシビスコアを高くし、上位で提示する。また、椎茸が嫌いなあるユーザに対しては、椎茸 (嫌いな食材) を多く含むレシピ A のレシビスコアを低くし、下位で提示する。その結果、ユーザの直感にあったレシピの順位付けが可能となる。

4.2 食材分量の分散による食材の位置付け

4.1 節にて、レシビスコアを算出する際に食材分量を考慮する必要性について述べた。しかしながら、食材の種類にかかわらず、単にレシピ中のある食材の分量が何 g 以上なら重みを大きくし、何 g 以下なら重みを小さくというような、単純なスコアリングを行うことは適切ではない。なぜなら、100g という食材分量が非常に多い場合もあれば、非常に少ない場合もある。したがって、全レシピ中における、ある食材の分量の平均値や分散を考慮する必要があると考えている。

図 4 に、全レシピ中における食材 k 、食材 p 、食材 q の分量の分布と平均値の例を示す。あるレシピにて利用される各食材の分量が 30g であったとしても、それぞれの意味が大きく異なることがわかる。

図 4 中の中段と下段に注目する。食材 p においては、30g という分量は非常に小さいことがわかる。したがって、このレシピにおいては食材 p の影響は小さいといえる。一方、食材 q においては、30g という分量は大きいことが分かる。したがって、このレシピにおいては、食材 q の影響は大きいといえる。したがって、各食材の分量の平均値を考慮する必要があることが分かる。

次に、図 4 中の上段と中段に注目する。食材 k も食材 p も分

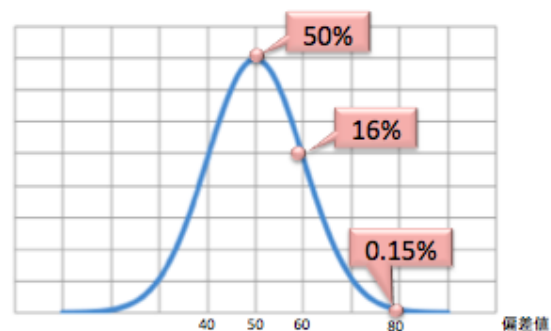


図 5 正規分布

量の平均値は同じであるが、その分散の状況に応じて、30g という分量の意味が異なることが理解できる。したがって、各食材の分量の分散についても考慮する必要があるといえる。

そこで、本手法では、レシピデータベース内の全レシピにおける、食材分量の分散を用いて各食材の位置付けを導出し、レシビスコアの算出に用いる。

4.3 食材分量を考慮した食材の重み付け

食材 k の含有量の平均および標準偏差を用いて、あるレシピにおける食材 k の重みを算出する。ただし、食材 k の食材分量はすべて 1 人分の分量に統一して扱う。平均はレシピデータベース内の全レシピ中、食材 k が含まれるレシピを抽出し、それらのレシピにおける食材 k の食材分量の平均を算出する。

さらに、標準偏差は上記で算出した平均を用いて、以下の式により算出する。

$$\sigma_g = \sqrt{\frac{1}{n} \sum_{i=1}^n (g_i - g_m)^2} \quad (6)$$

ただし、 n は食材 k を含むレシピ数、 g_i はレシピ i における 1 人分の食材 k の分量、 g_m は g_i の平均を示す。

重みを以下に付与する。

$$Score(R) = \sum_{k \in R} I_k W_k \quad (7)$$

W_k : 対象レシピにおける食材 k の重み

求められた偏差値を図 5 の正規分布を使って考える。偏差値 50 の食材は全体の上位 50 % を占めていると考える。偏差値 60 は全体の上位 16 %、偏差値 80 は全体の上位 0.15 % を占めているということになる。そしてこの正規分布全体に 0~2 の重みを割り当てる。偏差値が高ければ高いほど好き、嫌いの重みがそれぞれ大きくなり、逆に偏差値が引くければ低いほど好き、嫌いの重みがそれぞれ小さくなる。例えば偏差値が 60 を例とする。偏差値 60 の食材は全体の上位 16 % ある。仮にこの正規分布を 200 分割し、偏差値が低いものから順に 200 位までの順位を設ける。よって偏差値 60 ならば 168 位になる。これを 0~2 の重みに割り当て直して式 (7) の W_k のような形で付与してレシピ推薦スコアへ反映する。偏差値が 50 であれば 1.0 の重みとなり、偏差値 60 ならば 1.68 となる。分量を量らない食材 (豆板醤やバニラエッセンスなど) は少量でもレシピに大きく影

表 1 好きな食材として評価実験に用いた食材とレシピ

対象食材	対象食材の分量が多いレシピ	対象食材の分量が平均的なレシピ	対象食材の分量が少ないレシピ
鮭	鮭大根	鮭の照り焼き	塩鮭とネギの雑穀チャーハン
玉ねぎ	丸ごとオニオンスープ	肉野菜炒め	ベーコンと卵のピラフ
たらこ	和風たらこスパゲッティ	じゃがいもとたらこの炒め物	たらこのチーズ入り卵焼き

表 2 嫌いな食材として評価実験に用いた食材とレシピ

対象食材	対象食材の分量が多いレシピ	対象食材の分量が平均的なレシピ	対象食材の分量が少ないレシピ
セロリ	和風ポトフ	ラタトゥイユ	ペンネのミートソース
納豆	納豆チャーハン	えのき納豆かけ押し麦ご飯	イワシの納豆和え

響を与えるものと考え、レシピに含まれているかどうかで重み付けを行う。

また、ヒストグラムからもレシビスコアを算出することも可能である。こちらの方が実データに基づいて重み付けを行えるため今後検討を行う予定である。

5. 予備実験

提案手法の妥当性を確認するため、被験者実験を通じて以下の検証を行った。

5.1 実験方法

対象とするレシピは、味の素のレシピ大百科 [3] に登録されているレシピの中から我々が主菜と判断したレシピである。まず、ある食材に対して「大好き」から「大嫌い」までの 6 段階でアンケートを実施し、そのうち多くの被験者が大嫌い、嫌い、もしくは、好き、大好きと回答した食材を対象とした。表 1 に、好きな食材として評価実験に用いた食材とレシピを、表 2 に、嫌いな食材として評価実験に用いた食材とレシピを示す。表 1、表 2 に示すとおり、対象となった食材の分量が多いもの、平均的なもの、少ないもの、の 3 種類用意した。

被験者には、対象食材ごとに用意された 3 つのレシピの中から最も食べたいと思うレシピを 1 つ選択させると共に、選択した理由を尋ねた。なお、被験者は 20 代の男女計 7 名であった。

5.2 食材に対する嗜好推定手法の検証

5.2.1 好きな食材に対するレシピの選択結果

アンケートにて多くの被験者が好きと回答した食材は、鮭、玉ねぎ、たらこであった。実験結果は表 3 の通りである。

表 3 好きな食材を含むレシピのうち選択されたレシピの内訳

好きな食材の分量が多いレシピ	31%
好きな食材の分量が平均的なレシピ	44%
好きな食材の分量が少ないレシピ	25%

若干ではあるが、好きな食材の分量が少ないレシピよりは、好きな食材の分量が多いレシピを選ぶ割合が多かった。なお、平均的な分量のレシピを選んだ被験者の中にも、選んだ理由として“対象食材がどれだけ多く含まれているかを重視して選んだ”と答えた被験者もあり、好きな食材の分量についてはその分量がレシピの評価に影響があるということがわかった。

5.2.2 嫌いな食材に対するレシピの選択結果

アンケートにて多くの被験者が嫌いと回答した食材はセロリ、

表 4 嫌い食材を含むレシピのうち選択されたレシピの内訳

嫌いな食材の分量が多いレシピ	20%
嫌いな食材の分量が平均的なレシピ	20%
嫌いな食材の分量が少ないレシピ	60%

納豆であった。実験結果は表 4 の通りである。

嫌いな食材に関しては、明らかにその分量が少ないレシピを選択する割合が多かった。したがって、嫌いな食材の分量についても、その分量がレシピの評価に影響があるということが分かった。

6. おわりに

本稿では、好き嫌いラベルが付与された食材分量のレシビスコアへの影響について考察すると共に、新たなレシビスコア算出手法について検討した。予備実験の結果より、好きな食材や嫌いな食材に関しては、その分量がレシビスコアに影響があることを確認した。

今後の課題として推薦システムの実装を行い、提案手法の有効性の確認と改良を行う予定である。また調理法、レシピジャンルによる好き嫌い判定に基づくレシビスコア算出の検討も行う予定である。

謝 辞

本研究の一部は、京都大学グローバル COE プログラム「知識循環社会のための情報学教育研究拠点」(研究代表者：田中克己)、および文部科学省科学研究費補助金 基盤研究 (C)(研究代表者：中島伸介、課題番号 23500140) による。ここに記して感謝します。

文 献

- [1] 李福実, 上田真由美, 平野靖, 梶田将司, 間瀬健二. 個人の嗜好を考慮した料理レシピ推薦システムにおける栄養情報の取り扱いに関する検討, DEIM Forum 2009, E5-3, 2009.
- [2] 上田真由美, 高畑麻理, 中島伸介. レシピ閲覧・摂食履歴を用いた嗜好の抽出. Web とデータベースに関するフォーラム (WebDB Forum 2011), 情報処理学会シンポジウムシリーズ, 3G-1-2, 2011.
- [3] レシピ大百科, 味の素株式会社, <http://ajinomoto.co.jp/recipe/>
- [4] マギーレシピサイト, ネスレ日本株式会社, <http://www.recipe.nestle.co.jp/>
- [5] クックパッド, クックパッド株式会社, <http://cookpad.com/>
- [6] 楽天レシピ, 楽天株式会社, <http://recipe.rakuten.co.jp/>

- [7] Yahoo!レシピ, Yahoo! JAPAN,
<http://recipe.gourmet.yahoo.co.jp/>
- [8] Google レシピ検索, Google Inc.,
<http://www.google.co.jp/>
- [9] ユーザーサポート, クックパッド株式会社,
[http://info.cookpad.com/jobs/position /user-support](http://info.cookpad.com/jobs/position/user-support)
- [10] 三野陽子, 小林一郎. ユーザのスケジュールを考慮したダイエットのためのレシピ推薦, DEIM Forum 2009, A2-1, 2009.
- [11] 苺米志帆乃, 藤井敦. 栄養素等摂取バランスの分析に基づく食生活支援システム, 日本データベース学会論文誌, Vol.8, No.4, pp.1-6, 2010.
- [12] Jill Freye and Shlomo Berkovsky. Intelligent Food Planning: Personalized Recipe Recommendation. Proceedings of the 14th International Conference on Intelligent User Interfaces, 2010.
- [13] 塩澤秀和. 食材の優先度を考慮したビジュアルな料理レシピ検索インタフェース, 電子情報通信学会論文誌, Volume J94-A No.7, pp.458-466, 2011.