

レシピ推薦のためのユーザの食材購入周期に基づく 余剰食材・購入予定食材推定手法

吉田 朋史[†] 北山 大輔^{††}

[†] 工学院大学大学院工学研究科情報学専攻 〒163-8677 東京都新宿区西新宿1丁目24番2号

^{††} 工学院大学情報学部コンピュータ科学科 〒163-8677 東京都新宿区西新宿1丁目24番2号

E-mail: [†]em15020@ns.kogakuin.ac.jp, ^{††}kitayama@cc.kogakuin.ac.jp

あらまし 食料品を購入するユーザは多くの場合自宅等に余っている食材（余剰食材）を保有しており、これらの余剰食材の種類・量によってユーザが考える料理のレシピや購入する食材が変化するなど、ユーザの購入行動やレシピ選択に影響を与えていると考えられる。そこで、ユーザが定期的に購入している食材の購入周期を過去の購入履歴に基づき算出し、直近の購入日からの経過日数によって食材の余剰確率と購入確率を変化させることでユーザの余剰食材および購入予定食材を推測する手法を提案する。スーパーマーケットチェーンでのPOSデータをユーザの購買履歴として評価実験を行い、実際の商品購入活動を精度良く推定できているかを検証する。

キーワード POS データ, パーソナライゼーション, レシピ推薦

1. はじめに

スーパーマーケットなどで食品を購入する際、普段から周期的に一定の間隔で購入し続けている食材（周期的購入食材）を購入する際には、複数回買い物に行く手間や単価を考慮して1回の料理で使い切れる量よりも多い食材を購入することも少なくない。その結果、翌日以降もこれらの余った食材（余剰食材）が冷蔵庫等に備蓄されている場合が考えられる。これらの余剰食材は購入してからの経過日数が短い程充分な量が余っている可能性が高く、購入から日数が経過し余剰した分の食材を消費していくにしたがって、ユーザは新たにその食材を買い足す必要性を考慮しながら買い物を行うと考えられる。

また、料理のレシピを考える際にも余剰食材がユーザの選択に影響を与えると考えられる。例として、肉・野菜等の生鮮食品が余剰食材であった場合、食品の賞味期限は比較的短期であることが多いため、なるべく早く余剰食材を使いきれいなようなレシピをユーザは求めると考えられる。さらに、レシピを考えながら買い物をしているユーザに対して、余剰食材を含みかつもうすぐ購入する予定だった周期的購入食材（購入予定食材）を含むレシピを推薦することは、ユーザは既にある食材と普段から頻繁に使う食材だけで作れるレシピを知ることになり、有用なレシピ推薦となりうると考えられる。

マルハニチロホールディングスの「料理レシピに関する調査」[1]では、料理レシピを選ぶ際の基準として、54.0%の人が冷蔵庫にある食材にマッチするかどうかを基準しており、料理レシピを探すタイミングとしては、56.5%の人が作りたい料理が決まる前と回答している。このように、冷蔵庫の中にある余剰食材を推定することと購入前に購入予定食材を推定することはユーザの購入行動やレシピ選択に大きく影響を与えると考えられる。

ユーザの余剰食材を把握する機能の例として、家電大手の

シャープ（株）は音声認識機能などを搭載した冷蔵庫を販売している[2]。この冷蔵庫は、冷蔵庫に食材を収納する際や冷蔵庫から食材を取り出す際にユーザが食品名を声で入力することで、冷蔵庫内に収納されている食材の種類・量を把握し、ユーザが買い足す可能性がある食材や賞味期限が近い食材を使ったレシピなどをユーザに推薦する機能を持っている。

このような既存の実用例では、食材を追加・使用するたびにユーザが何らかの方法で食材をシステムに登録する必要があり、ユーザがシステムを使用する際の負担や手間が大きいことが課題として考えられる。そこで、本研究では過去のユーザの食材購入履歴からユーザの周期的購入食材の購入周期を算出し、購入からの経過日数によって食材の余剰している確率を変化させることで、ユーザの余剰食材を推測する手法を提案する。スーパーマーケットなどが保有するユーザのPOSデータなどをユーザの食材購入履歴としてもちいることで、ユーザがシステムを使用するための導入コストが軽減されることが期待される。

また、提案手法によって推測したユーザごとの余剰食材と購入予定食材をもちいることで、ユーザ個々に最適化された食材推薦・レシピ推薦が行えると考えられる。本稿では、余剰食材の推定に関して被験者を持ちいた評価を行う。

以下に本稿の概要を記す。2節では関連研究について、3節では提案手法である食材の周期的購入性の算出と各食材の余剰確率・購入確率の算出について、4節では実際のスーパーマーケットチェーンのPOSデータをもちいた提案手法の評価実験について、5節ではまとめと今後の課題を述べる。

2. 関連研究

余剰食材を効率的に使うことができるレシピ推薦手法についての研究では、文献[3][4][5][6]がある。[3]ではユーザがこれまでに作ったレシピを基にレシピに含まれる食材の消費量を推測し、余剰食材を無駄なく使い切ることができるレシピを推薦

する手法を提案している.[4]では冷蔵庫の食材の分量や賞味期限を考慮したレシランキング手法をもちいて、冷蔵庫に余っている食材を効率良く使い切るためのレシピ検索システムを提案している。

[5]では嗜好・栄養・カロリー・調理の簡単さ・在庫食材情報・同じレシピの連続推薦の回避の6つを主なレシピ決定の要因であると想定し、これらの要因の数値化の方法と、各要因の数値にかける重みをユーザの価値観や生活スタイルに合わせて最適化する方法を提案する.[6]では調理履歴から算出した嗜好食材に基づく推薦、余剰食材を考慮した推薦、調理履歴のレシピの調理法の3点を考慮し、食材の重み付けを変化させることで利用者個人の嗜好に合わせたレシピ推薦を行う。

これらの研究ではユーザは購入した食材や冷蔵庫に余っている食材の種類・分量をデータベースに登録する必要があるが、本研究では余剰食材そのものを推定することを目的とすることや、店舗が持つユーザの購買履歴のみをデータとして用いるため、ユーザが事前に必要な情報を登録する必要が無い点で異なる。

既存の調理履歴や使用食材履歴をもちいたレシピ推薦を行う研究としては[7][8]がある.[7]ではユーザが調理できる料理のレパートリーを拡大することを目的として、ユーザの調理履歴からレシピ経験値と食材経験値を含む調理経験値を推定し、調理経験値に応じたレシピを推薦する。また、操作画面からレシピ経験値と食材経験値を操作することで、調理経験のある食材を増やしたりより複雑な調理を必要とするレシピに挑戦するなど、ユーザは自分の要求に基づいてインタラクティブにレシピを検索できる。

[8]では文書検索における単語の特徴度の尺度である TF-IDF(Term Frequency - Inverted Document Frequency)を食材の特徴度に応用した FF-IRF (Foodstuff Frequency - InvertedRecipe Frequency)を提案し、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したレシピ推薦システムを提案している。提案手法では、レシピ毎の食材の使用量や過去にユーザが調理したレシピなどの情報はもちいず、特定の日時における余剰食材と購入予定食材を推定することで将来的にユーザにとって満足度の高いレシピ推薦を行うことを目指している。また、ユーザにとってより直感的で理解しやすいレシピ検索システムについての研究では[9][10][11][12]などがあげられる。

3. ユーザの購入周期に基づく余剰食材推定手法

本研究では、ユーザの過去の食材購入履歴に基づき、指定された日付におけるユーザの余剰食材と購入予定食材を推定する手法を提案する。提案手法では、まずユーザが購入したことのある食材の周期的購入性を購入間隔日数の中央値とその標準偏差をもちいて算出したうえで、指定された日付における余剰確率と購入確率を推定する。3.1節では食材の周期的購入性の算出方法について、3.2節では余剰確率と購入確率推定手法について述べる。

3.1 食材の周期的購入性の算出

各食材がどの程度周期的に購入されるかは、購入するユーザや食材自身の性質によって大きく左右されると考えられる。本研究では食材が一定範囲内の購入間隔日数で周期的に購入されることを前提として食材の余剰確率と購入確率を算出することを目的としているため、周期的に購入されていない食材に対しては算出された余剰確率と購入確率では推定精度が減少すると考えられる。そこで、算出された余剰確率と購入確率の信頼性を測る指標として式(1)を考える。

$$Periodicity(\bar{t}, \sigma_o) = \frac{\sigma_o}{\log_{10} \bar{t}} \quad (1)$$

式(1)では、ユーザが該当食材を購入してから次回購入するまでに経過した日数を購入間隔日数と定義し、ユーザの該当食材に対する全購入履歴に対する購入間隔日数の中央値を \bar{t} 、購入間隔日数の標準偏差を σ_o としている。標準偏差 σ_o は購入間隔日数の中央値 \bar{t} の統計的なバラツキの具合を表すため、 \bar{t} に対して σ_o が占める割合が小さいほどユーザがその食材を一定の周期で購入していると考えられるため、式(1)の値が小さい食材ほど購入間隔日数にバラツキがなく定期的に購入されており、算出した余剰確率・購入確率の値の信頼性が高くなると考えられる。

3.2 余剰確率と購入確率の算出

ユーザが購入したことのある各食材に対して、前回の購入日から指定された日付までの経過日数 t を変数とする食材ごとの余剰確率関数と購入確率関数を算出することで、指定された日付におけるユーザの余剰食材と購入予定食材を推定する。各関数の算出には、ユーザの食材購入履歴から算出した購入間隔日数の中央値 \bar{t} と標準偏差 σ_o をもちいる。

3.2.1 余剰確率

余剰食材は購入してからの経過日数が短いほど充分な量が余っている可能性が高く、経過日数が増加するごとに徐々に使用されていくと考えられる。そこで、提案手法では余剰確率関数が以下の条件を満たすとする。

- (1) 直近の食材購入日を余剰確率関数の始点とし、食材購入日での余剰確率を最大値1とする
- (2) 余剰確率の減少は食材購入日を中心とする正規分布に従う
- (3) 購入日から $\bar{t} + \sigma_o$ 日後における余剰確率は5%となる

条件(1)・(2)に基づき、ユーザが該当食材を購入した直近の日付から t 日後における該当食材の余剰確率は式(2)で定義される。

$$P_s(t) = 1 - 2 \cdot \frac{1}{\sqrt{2\pi}\sigma_s} \int_0^t \exp\left(-\frac{t}{2\sigma_s^2}\right) dt \quad (2)$$

このとき、式(2)における正規分布の分散 σ_s は、条件(3)を満たすよう式(3)により決定する。

$$1 - 2 \cdot \frac{1}{\sqrt{2\pi}\sigma_s} \int_0^{\bar{t} + \sigma_o} \exp\left(-\frac{t}{2\sigma_s^2}\right) dt = 0.05 \quad (3)$$

表 1 実験に用いたユーザ

	ユーザ A	ユーザ B
購入食材カテゴリ数	198	226
購入食材カテゴリ数	3734	7600

表 2 ユーザ A の Periodicity が低い 5 食材

	納豆	オイスターソース	肉まん	豚切り落とし	木綿豆腐
Periodicity	5.20	5.47	5.54	6.32	6.57
総購入数	137	4	3	259	92
購入間隔日数の中央値 \bar{t}	5.0	103.0	34.0	4.0	7.5
購入間隔日数の標準偏差 σ_s	3.64	11.02	8.49	3.81	5.75

表 3 ユーザ A の Periodicity が高い 5 食材

	とうもろこし	唐揚げ菜	あずきパー	すいか	冬瓜
Periodicity	270.91	193.65	158.41	154.08	138.89
総購入数	4	7	4	6	4
購入間隔日数の中央値 \bar{t}	4.0	12.0	62.0	10.0	34.0
購入間隔日数の標準偏差 σ_s	163.10	208.98	283.93	154.08	212.70

3.2.2 購入確率

食材を購入する確率は、購入してからの経過日数が短いほど新たに購入する確率が低く、経過日数が増加するごとに徐々に購入する確率が増加していくと考えられる。また、直近の購入日から平均購入間隔日数 \bar{t} 日が経過した日付が食材を購入する確率が高いと考えられる。そこで、提案手法では購入確率関数が以下の条件を満たすとす。

(1) 直近の食材購入日より $\bar{t} - \sigma_s$ 日後を購入確率関数の始点とする

(2) 購入確率は食材購入日から \bar{t} 日後を中心とする正規分布に従う

(3) 購入日から \bar{t} 日後における購入確率は 95% となる

条件 (1) と条件 (2) に基づき、ユーザが該当食材を購入した直近の日付から t 日後における該当食材の購入確率は式 (4) で定義される。

$$P_p(t) = 2 \cdot \frac{1}{\sqrt{2\pi}\sigma_p} \int_{\bar{t}-\sigma_s}^t \exp\left(-\frac{t-\bar{t}}{2\sigma_p^2}\right) dt \quad (4)$$

式 (4) における正規分布の分散 σ_p は、条件 (3) を満たすよう式 (5) により決定する。

$$2 \cdot \frac{1}{\sqrt{2\pi}\sigma_p} \int_{\bar{t}-\sigma_s}^{\bar{t}} \exp\left(-\frac{t-\bar{t}}{2\sigma_p^2}\right) dt = 0.95 \quad (5)$$

4. 実験と考察

経営科学系研究部会連合協議会主催、平成 27 年度データ解析コンペティションにおいて提供された (株) アイディーズの i-code データをもちいて、実際の商品購買履歴に対して提案手法を適用し特定の日時での余剰確率と購入確率を算出する。実験には、商品購買履歴から無作為に抽出したユーザ A・B の 2 名のデータ (表 1) をもちいた。

4.1 食材の周期性の算出

ユーザ A・B の購入履歴から購入の周期性を表す Periodicity を算出し、ユーザ A・B ごとに Periodicity が小さい順に 5 食材を抽出した結果を表 2・3 に、Periodicity が大きい順に 5 食材を抽出した結果を表 4・5 に示す。

表 4 ユーザ B の Periodicity が低い 5 食材

	フライドチキン	ドレッシング	きのこ	納豆	木綿豆腐
Periodicity	0.45	0.852	2.70	2.73	2.81
総購入数	3	3	3	256	244
購入間隔日数の中央値 \bar{t}	34.5	309.5	20.5	3.0	3.0
購入間隔日数の標準偏差 σ_s	0.71	2.12	3.53	1.30	1.34

表 5 ユーザ B の Periodicity が高い 5 食材

	いかなご	ぶどう	梨	焼きそば惣菜	サブ味噌缶
Periodicity	225.81	200.34	174.19	151.79	150.19
総購入数	4	5	13	4	3
購入間隔日数の中央値 \bar{t}	22.0	10.0	4.0	51.0	306.0
購入間隔日数の標準偏差 σ_s	303.14	200.34	104.87	259.19	373.35

表 2・4 より、Periodicity が低い食材として「納豆」「木綿豆腐」など総購入数も多く購入間隔日数の標準偏差も小さい食材の他に、「オイスターソース」「ドレッシング」など総購入数が少なく購入間隔日数の中央値も大きいが、購入間隔日数の標準偏差は小さい食材が抽出されていることがわかる。調味料やドレッシングなどの食品は、一度購入すると使いきるまでに数ヶ月かかることが多いため頻繁に購入するわけではないが、毎日の食事に定期的に使用され、使いきれば買い足されていく場合が多いと考えられるため、周期的購入食材として適切であると考える。

また、表 3・5 より Periodicity が高い食材として「すいか」「梨」「冬瓜」など、特定の季節に旬を持つ食材が多く含まれていることがわかる。これらの食材では購入間隔日数の中央値 \bar{t} に対して購入間隔日数の標準偏差 σ_s の値が非常に大きくなる傾向が見られるが、その原因は提案手法での σ_s の算出方法にあると考えられる。提案手法では食材ごとの σ_s の算出に全購入履歴の購入間隔日数をもちいているため、履歴内の他の購入間隔日数に対して突出して長い購入間隔日数が少数でも含まれた場合、 σ_s の値が大きく変化してしまうことが考えられる。

特定の季節に旬を持つ食材は、旬の季節が終わると次の旬の季節が訪れるまで購入する機会が減少し、外れ値となる突出して長い購入間隔日数が含まれやすい傾向にあるために標準偏差が大きくなりやすいと考えられる。また、季節の影響を受けにくい食材であっても、旅行や帰省などで長期間買い物を行わなかった場合に大きく購入間隔が空いてしまうことが考えられる。そのため、このような全体の傾向と異なる外れ値を考慮して σ_s を算出する必要があると考えられる。

また、表 2・4 より Periodicity が低い 5 食材における各食材の Periodicity の値はユーザ B の方がユーザ A に比べて小さい傾向にあることがわかり、ユーザや食材によって Periodicity の値が異なることが確認できた。

4.2 余剰確率と購入確率

表 1 のユーザ A・B に共通する周期的購入食材である「豚切り落とし」「小松菜」「卵」の 3 食材の 2014 年 1 月 28 日における各食材の余剰確率と購入確率を表 6・7 に示す。

表 6 より、ユーザ A の 2014 年 1 月 28 日における「豚切り落とし」の購入確率は 0.928 と非常に高い値を示しており、実際にその翌日の 1 月 29 日にユーザ A は「豚切り落とし」を購入していることがわかる。また、「豚切り落とし」の前日に購

表 6 ユーザ A における各食材の余剰確率と購入確率

	豚切り落とし	小松菜	卵
購入間隔日数の中央値 \bar{t}	4.0	14.0	4.0
購入間隔日数の標準偏差 σ_o	3.806	9.645	4.057
余剰確率関数の標準偏差 σ_s	3.983	18.126	4.455
購入確率関数の標準偏差 σ_p	1.950	10.952	2.072
各食材の直近購入日	1月22日	1月21日	1月25日
直近購入日からの経過日数 t	6	7	3
1月28日時点での余剰確率 $P_s(t)$	0.132	0.699	0.466
1月28日時点での購入確率 $P_p(t)$	0.950	0.472	0.579
実際の各食材の次回購入日	1月29日	2月4日	1月30日

表 7 ユーザ B における各食材の余剰確率と購入確率

	豚切り落とし	小松菜	卵
購入間隔日数の中央値 \bar{t}	5.0	5.0	11.5
購入間隔日数の標準偏差 σ_o	7.513	6.198	6.037
余剰確率関数の標準偏差 σ_s	7.727	6.645	9.532
購入確率関数の標準偏差 σ_p	3.834	3.166	3.081
各食材の直近購入日	1月24日	1月28日	1月24日
直近購入日からの経過日数 t	4	4	4
1月28日時点での余剰確率 $P_s(t)$	0.533	1.000	0.655
1月28日時点での購入確率 $P_p(t)$	0.744	0.064	0.036
実際の各食材の次回購入日	2月13日	1月28日	2月11日

表 8 ユーザ A の「豚切り落とし」購入確率の経過日数による変化

直近購入日からの経過日数 t	2	3	4
購入間隔日数の中央値 \bar{t}	4.0	4.0	4.0
購入間隔日数の標準偏差 σ_o	3.806	3.806	3.806
各食材の直近購入日	4月20日	4月8日	4月11日
実際の購入日	4月22日	4月11日	4月15日
実際の購入日での余剰確率 $P_s(t)$	0.254	0.557	0.949

表 9 ユーザ A の「小松菜」購入確率の経過日数による変化

直近購入日からの経過日数 t	9	13	14
購入間隔日数の中央値 \bar{t}	14.0	14.0	14.0
購入間隔日数の標準偏差 σ_o	21.374	21.374	21.374
各食材の直近購入日	3月17日	9月2日	8月18日
実際の購入日	3月26日	9月15日	9月2日
実際の購入日での余剰確率 $P_s(t)$	0.597	0.876	0.949

入している「小松菜」に関しては、直近購入日からの経過日数は「豚切り落とし」と1日しか変わらない7日となっているが購入確率は0.472と「豚切り落とし」に比べて低くなっており、実際に次回購入日は2月4日と「豚切り落とし」に比べ購入間隔日数に差があるなど、提案手法で推定した余剰確率と購入確率が実際の購入履歴データとうまく一致していることが確認できた。

また、表6・7に示した購入確率の日数毎の変化に注目すると、購入間隔日数の中央値 \bar{t} と標準偏差 σ_o が小さい「豚切り落とし」では、経過日数 t が1日異なることによる購入確率の変化が「小松菜」よりも大きいことがわかる。提案手法では購入間隔日数の中央値を中心とする正規分布を用いて購入確率を推定するため、 \bar{t} や σ_o が小さいほど購入確率の変化が大きくなりやすく、その結果「小松菜」での実際の購入日での購入確率の平均値が「小松菜」「卵」に比べて低下したと考えられる。

4.3 レシピ推薦への応用

表4・5より、1月28日時点で余剰確率が50%を超えている食材としてユーザAでは「小松菜」「卵」、ユーザBでは「豚切り落とし」「小松菜」「卵」があげられる。これらの余剰確率が

高い食材を含むレシピとして、「厚揚げと小松菜の卵炒め」^(注1) や「小松菜とシーチキンのチヂミ」^(注2) などが考えられ、レシピに含まれる余剰確率の低い食材や購入確率の高い食材と共にこれらのレシピをユーザに提示することで効果的なレシピ推薦や購入予定食材推薦が行えると考えられる。

5. まとめと今後の課題

本研究では、ユーザが定期的に購入している食材の購入周期を過去の購入履歴に基づき算出し、購入からの経過日数によって食材の余剰確率と購入確率を変化させることでユーザの余剰食材を推測する手法を提案した。スーパーマーケットチェーンのPOSデータをもちいて実際の購入履歴データに対して提案手法を適用する実験を行い、周期的購入食材の抽出や比較的購入周期が長い食材に対して提案手法の有効性を確認した。

今後の課題として、購入周期が短い食材における推定精度の改善や、算出した食材の周期的購入性を考慮するよう余剰確率・購入確率の算出方法を改善していくことなどが挙げられる。さらに、食材ごとの賞味期限や消費のされ方の違いなど、食材ごとの特性を考慮していくことなどを検討している。

謝 辞

本研究の一部は、平成27年度科研費若手研究(B)(課題番号:15K16091)によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 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. https://www.maruha-nichiro.co.jp/news_center/research/pdf/20130227_recipe_cyouusa.pdf.
- [2] シャープ. 冷蔵庫 | ともだち家電: シャープ. <http://www.sharp.co.jp/tomodachi/products/reizo/>.
- [3] 木原ひかり, 上田真由美, 中島伸介. 余剰食材の使い切りを考慮したレシピ推薦手法の提案. 第3回 DEIM Forum, pp. E3-3, 2011.
- [4] 赤澤康幸, 宮森恒. 冷蔵庫食材を考慮した料理レシピ検索システムの提案. 第4回 DEIM Forum, pp. E3-3, 2012.
- [5] 伊原啓晃, 玉井森彦, 安本慶一. 複数要因を総合的に考慮した夕食レシピ推薦システムとその評価. 研究報告コンピュータセキュリティ (CSEC), Vol. 2014, No. 32, pp. 1-8, 2014.
- [6] 望月美里, 高橋裕樹. 余剰食材の有効活用と嗜好に基づく多彩なレシピ推薦. *ITE Technical Report*, Vol. 36, No. 8, pp. 127-130, 2012.
- [7] 義貴中岡, 哲司佐藤. レパートリー拡大のための戦略的レシピ推薦システムの実装と評価 (データ工学と食メディア). 電子情報通信学会技術研究報告. DE, データ工学, Vol. 113, No. 214, pp. 25-30, sep 2013.
- [8] 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. *DBSJ letters*, Vol. 6, No. 4, pp. 29-32, 2008.
- [9] 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 A, Vol. 94, No. 7, pp. 458-466, 2011.
- [10] 塩澤秀和. キーワードとその優先度による対話的な料理レシピ検

(注1): <http://cookpad.com/pro/recipes/2758882>

(注2): <http://cookpad.com/recipe/3548165>

索. *WISS 2007*, 2007.

- [11] 野間田佑也, 星野准一. Graphicalrecipes : レシピ探索支援のための視覚化システム. 芸術情報学会論文誌, Vol. 7, No. 2, pp. 43-54, 2008.
- [12] 苅米志帆乃, 藤井敦. 栄養素等摂取バランスを考慮した料理レシピ検索システム (コンテンツ技術,web 情報システム). 電子情報通信学会論文誌. D, 情報・システム, Vol. 92, No. 7, pp. 975-983, 2009.