

# 地域度を考慮したメニューの推薦法に関する研究

佐野 駿<sup>†,††</sup> 佐藤 哲司<sup>††</sup>

<sup>†</sup> 筑波大学情報学群知識情報・図書館学類 〒305-8550 茨城県つくば市春日 1-2

<sup>††</sup> 筑波大学大学院図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

E-mail: <sup>†</sup>{shuntasato,satoh}@ce.slis.tsukuba.ac.jp

**あらまし** 近年、グルメサイトの認知度の向上は目覚ましく、多くの利用者が旅行や非日常的な出来事で外出する際に利用している。店舗の利用者が記入したレビュー記事数やメニューの推薦者数は、そのメニューの話題性を表し、メニューの評価値は、利用者の満足度を表していると考えられ、これらの指標を用いて、利用者は店舗やメニューを選択している。本論文では、メニューの話題性や満足度に加えて、その土地独自で限定的なメニューであることを表す地域性を考慮したメニューの推薦法を提案する。提案法では、メニューに対するレビュー記事数や推薦者数は、そのメニューの話題性を表す指標、メニューの評価値は、利用者の満足度を表わす指標と考え、これらの指標と地域性から、地域度を算出する。実際に運用されている商用レビューサイトから収集した情報に基づいて、いくつかのメニューに対して地域度を算出した結果、その土地独自のメニューを推薦できることを確認した。

**キーワード** 地域度、メニュー推薦、ビッグデータ

## A recommendation method on restaurant menu by considering locality

Shun SANO<sup>†,††</sup> and Tetsuji SATOH<sup>††</sup>

<sup>†</sup> College of Knowledge and Library Science, School of Informatics, University of Tsukuba

1-2 Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

<sup>††</sup> Graduate School of Library, Information and Media Studies, University of Tsukuba

1-2 Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

E-mail: <sup>†</sup>{shuntasato,satoh}@ce.slis.tsukuba.ac.jp

**Abstract** Recently, Gourmet-Sites have increased and they are used when users go to events. The number of review articles and the number of recommending menu written by store users express the level of Topic. And menu's evaluation points express the level of user's Satisfaction. Users eat the restaurant menu by using these factors. And I define local and particular restaurant menu as Local Restriction. In this paper, I propose the way of recommendation method on restaurant menu considering Local Restriction in addition to the Topic and Satisfaction. I define the three words, Local Restriction, Topic, Satisfaction as Locality, and I make a study of the way to recommend menu considering Locality.

**Key words** Chiikido, Recommend Menu, BigData

### 1. はじめに

近年、グルメサイトが増加してきており、このようなサイトは旅行や非日常的な出来事で外出する際に利用されている。特にサイトに記載されているレビュー記事は、店舗やメニューの選択の参考になっている。ぐるなび<sup>(注1)</sup>は、店舗の名前や所在地、メニューの一覧、メニューに対するレビュー記事、そのメニューを推薦する人数、そのメニューを評価する星数など多く

の情報が掲載されている。

店舗の利用者が記入したレビュー記事数やメニューの推薦者数は、そのメニューの話題性を表し、メニューに与えられた評価値は、利用者の満足度を表していると考えられる。多くの利用者はこれらの情報に基づいて、食事の際に店舗やメニューを選択していると考えられる。一方で旅先や外出先で食事をする際には、利用者はメニューの話題性や満足度に加えて、その土地ならではのメニューを選択すると考えられる。本論文ではこのような地域特有のメニューであることを地域性と呼ぶ。

本論文における地域性とは、その土地独自で、その土地での

(注1) : ぐるなび <http://www.gnavi.co.jp/>

み飲食できる限定的なメニューであることを測る指標である。例えば、山陰地方を旅行する利用者は、話題性や満足度でメニューを選択する。しかし、選択したメニューの中には、全国展開されているものも少なからず含まれている。本論文では、話題性や満足度だけでなく、このようなどこでも食べられるメニューではなく、その土地でなければ食べられない地域性の高いメニューの推薦を目指す。

本論文では、旅行や非日常的な出来事で外出をする際に、地域性・話題性・満足度の3つの視点からメニューを推薦することを目的とする。上記の3項目から算出される指標を地域度と呼ぶ。

地域度とは地域性・話題性・満足度の評価値を組み合わせた指標で、地域度が高ければ高いほど推薦する価値が高いメニューであることを意味する。メニュー選択には、メニューを様々な面から推薦する必要がある。旅行や非日常的な出来事で外出する際にメニューを選択すると限ったとき、地域性は特定のエリアやメニュー等その土地独自である希少性を測るもので、空間的局所性を示している。一方で話題性や満足度は、利用者がメニューの満足度や評価が高まりに感心しており、時間的局所性を示している。この空間・時間の両視点から、地域度を算出し、メニューの推薦を行う。

第2章では従来の研究を紹介し、本研究の位置付けを明確にする。第3章ではメニュー情報取得の方法、及びメニューの地域度を算出する方法を提案する。第4章では提案手法の評価を行い、第5章ではまとめを述べる。

## 2. 先行研究

グルメサイト関連の推薦に関する先行研究としては、飲食店等の「地域スポット推薦」、レビューからある情報を抽出する「特徴量抽出」をテーマにした研究がある。

地域スポット推薦をテーマにした研究には、奥ら[1]の研究がある。その地域でしか利用できない店舗等のスポットを、推薦する指標である地域限定性の算出を行っている。地域限定性算出のため、Web上から位置情報、テキスト情報を収集する。スポットの周囲で範囲を限定して、スポットの様々な観点から要素を抽出し、地域限定性を算出し、その有用性を示している。更に奥ら[2]は、地域限定性を考慮して、情報推薦を行っている。[1]の研究で抽出された語句の中には、ノイズとなる語句も多く含まれており、IDF値を限定性スコアに追加するなどしてノイズを除去し、地域限定性推薦の精度を高めている。手塚ら[3]は、ウェブコンテンツへのアプローチから地域性推定を行っている。検索手法としてクエリを投げてウェブ検索結果を解析し、フィルタリングなどを行うことで、クエリが関連する地域の取得を行っている。

また特徴量抽出をテーマにした研究は、小林ら[4]によって行なわれている。レビュー内容を自動的に分析し、製品の特徴をわかりやすく分析している。ユーザの意見が生まれた「理由」を製品の重要な特徴として扱い、形態素解析や係り受け解析によって、レビューの「理由」が書かれている部分の抽出を行っている。Gabrilovichら[5]は、Wikipediaに記載される概念か

らテキスト分類を行っている。テキストという自然言語間の類似度を測ることで、語と語の意味的関連性を見出している。

本論文ではメニュー単位で地域度を算出している。メニュー、店舗、地域の3つの視点から地域性、及び話題性や満足度を考慮した地域度を算出している。加えて、コンテキスト以外のメニュー・店舗・地域を考慮した地域性も推薦の一因としてしており、奥らの研究[1][2]とは異なると考えられる。また手塚らと共通である地域性をテーマにした研究であるが、ぐるなびのようなWeb上のビッグデータを扱ったテキスト処理・分析を行っており、手塚ら[3]と異なると考えられる。特徴量抽出に関する[4][5]との比較においても、本論文では地域度を算出し、メニューを推薦するという目的があるという相違点がある。本論文では、メニュー単位で評価して推薦することによって、ユーザに対してより詳細にメニュー推薦をすることができると考えられる。飲食店舗情報だけでなく、店舗ごとに記載されているメニュー情報も用いて考える。本論文はメニュー・店舗・地域から算出される地域性と、話題性や満足度を考慮した地域度という指標を算出し、メニューの話題性や満足度、地域性を考慮したメニュー推薦を行う点に新規性がある。

## 3. 地域度の算出法の提案

地域度は、1.節で述べたように、地域性・話題性・満足度を掛け合わせた指標で、評価値が高ければ高いほど推薦する価値が高くなるものとする。地域度算出の流れは、図1のようになっている。まずぐるなびからメニューデータを取得する。取得する全データをメニューDBと呼び、「メニュー」、「店舗」、「地域」、及び「レビュー」の各テーブルを作成する。このうち「メニュー」、「店舗」、及び「地域」は地域性を算出し、「レビュー」は話題性と満足度を算出する。地域性と話題性、及び満足度を算出した後、地域度を算出する。以下3.1節で、地域度を算出する為のメニューDBの取得方法を述べる。次に3.2節で、メニューの体系化について述べ、3.3節で、地域度の評価手法について述べる。

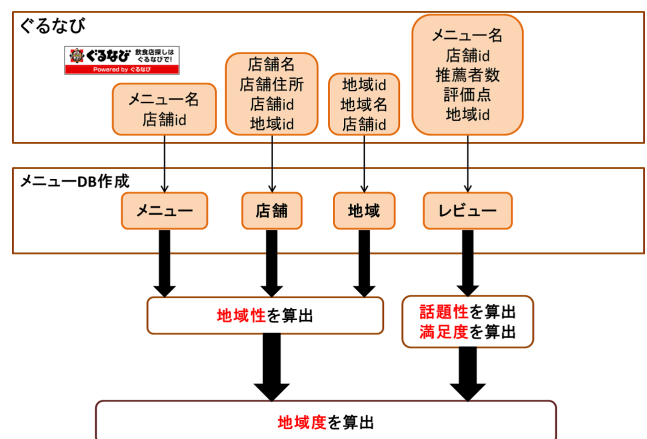


図1 提案手法の流れ

### 3.1 メニュー DB の取得方法

#### 3.1.1 評価データ

ぐるなびは月間約9億ページビュー、月間のぐるなびユニークユーザ数2,800万人、会員数は941万人を抱える大きなグルメレビューサイトである。ぐるなびはぐるなびWebサービス<sup>(注2)</sup>(以降からぐるなびAPIと呼ぶ)という、店舗の地域や都道府県、カテゴリなどの情報が公開されているサービスや、メニューページの充実など、メニューDBを網羅的に収集できると考えられる。ぐるなびAPIから都道府県ごとにレストラン情報を取得する。図2のように、店舗ごとに店舗ID、店舗名、店舗住所など多くの情報が掲載されている。こういった情報から店舗ごとに付与されている店舗IDを都道府県ごとにまとめる。

```
<?xml version="1.0" encoding="UTF-8"?>
<response>
  <total_hit_count>22106</total_hit_count>
  <hit_per_page>999</hit_per_page>
  <page_offset>1</page_offset>
  <pr_header>
    <id>g407400</id>
    <update_date>2012-10-15 12:38</update_date>
    <name>あひ庵 鳥安</name>
    <name_kana>アヒア 鳥安</name_kana>
    <latitude>35.689581</latitude>
    <longitude>139.790275</longitude>
    <category>会館・会館</category>
    <url>http://r.gnavi.co.jp/g407400/?ak=J4qfmr8oYzFC3JE5f5qJbCgpXecT%2F%2BuraTWzroeWo%3D</url>
    <url_mobile>http://mobile.gnavi.co.jp/shop/g407400/?ak=J4qfmr8oYzFC3JE5f5qJbCgpXecT%2F%2BuraTWzroeWo%3D</url_mobile>
    <image_url>
      <shop_image1>http://apicache.gnavi.co.jp/image/rest/index.php?img=g407400v.jpg&sid=g407400</shop_image1>
      <shop_image2>http://apicache.gnavi.co.jp/image/rest/index.php?img=g407400l.gif&sid=g407400</shop_image2>
      <qr_code>http://apicache.gnavi.co.jp/image/rest/index.php?img=g407400qr.png&sid=g407400</qr_code>
    </image_url>
    <address>〒103-0004 東京都中央区東日本橋2-11-7</address>
    <tel>03-3862-4008</tel>
    <fax>03-3862-4009</fax>
    <opentime>17:00~22:00</opentime>
    <holiday>日曜</holiday>
    <access>
      <line>地下鉄都営浅草線</line>
      <station>東日本橋駅</station>
      <station_exit></station_exit>
      <walk>5</walk>
      <note></note>
    </access>
    <pr>
      <pr_short>東日本橋は開田川沿いの創業明治五年の老舗合鴨料理屋。</pr_short>
      <pr_long>打ち子が揃った音ながらの石置の玄関が老舗料理屋の趣と風格を色濃く残している。料理はコースのみで他にメニューはなく音ながらの変わらぬ味を今に伝える名店として名高い。すべて個室で予約。皮付きのまま焼く切ったダキと呼ばれる鶏肉を鉄板焼きの要領でソイソイ焼きたり、あつちき焼いたたりき焼きたり焼きたり。</pr_long>
    </pr>
  </pr_header>
</response>
```

図2 ぐるなびAPIのレストラン情報のデータ構造

#### 3.1.2 メニュー DB の取得手順

店舗IDを都道府県ごとにまとめた上で、以下で各テーブルの取得方法を述べる。

##### 店舗テーブルの取得方法

(1) 店舗IDから得られる店舗URL「http://r.gnavi.co.jp/店舗ID/」を、都道府県ごとにcsvファイルに入れる。取得したcsvファイル毎に店舗URLを開き、店舗のホームページを取得する。

(2) 取得した店舗のホームページを開きながら、店舗テーブルカラムに必要な要素を正規表現で抽出する。正規表現は表1のようにパターンマッチさせる。

##### メニューテーブルの取得方法

(1) 店舗URLに「menu 数字.html」を加えて、メニューURL「http://r.gnavi.co.jp/店舗ID/menu 数字.html」を作成する。

(2) 数字には1ずつ入れていき、妥当性から3までとしてメニューのホームページを取得する。

(3) 取得したメニューのホームページを開きながら、メニューテーブルカラムに必要な要素を正規表現で抽出する。正規表現は表1のようにパターンマッチさせる。

##### 地域テーブルの取得方法

(1) ぐるなびの都道府県ごとのページにアクセスし、都道府県名を取得する。

##### レビューテーブルの取得方法

(1) レビューのホームページを見るとレビューURLは「http://mr.gnavi.co.jp/search/gid/店舗ID&p=数字」の形になっているので、店舗IDを利用して、レビューURLを作成する。

(2) 数字には1ずつ入れていき、妥当性から3までとしてレビューのホームページを取得する。

(3) 取得したレビューのホームページを開きながら、レビューテーブルカラムに必要な要素を正規表現で抽出する。正規表現は表1のようにパターンマッチさせる。

表1 メニューDB取得のための正規表現一覧

テーブル名	カラム名	正規表現
メニュー	メニュー	/(<dt class="itemTitle">)(.*)</dt>/s
メニュー	店舗ID	/(<meta name="shopid" content="(.*?)" \"/>)/s
店舗	店舗ID	/(<meta name="shopid" content="(.*?)" \"/>)/s
店舗	店舗	/<title>ぐるなび - (.*)</title>/s
店舗	店舗住所	/<td class="adr">(.*?)</td>/s
店舗	都道府県ID	/(<meta name="pref" content="(.*?)" \"/>)/s
レビュー	メニュー	/(<class="col_dred01">(.*?)</a></p>)/
レビュー	店舗ID	/(<a href="http://r.gnavi.co.jp/">(.*?)</a> class="col_brown01"/>)/s
レビュー	評価点	/(<alt=""/></span><strong>(.*?)</strong><span class="px12">点)/s
レビュー	推薦者数	/(<p><span class="px12">おすすめ人数</span><strong>(.*?)</strong><span class="px12">人)/s
レビュー	投稿件数	/(<p class="kurikomi_list_data">投稿件数:<span>(.*?)</span>)/s

#### 3.1.3 収集したメニューの分析

3.1.2で取得したメニューDBの各テーブルの取得件数は、メニューテーブルが約297万件、店舗テーブルが約6万件、地域テーブルが47件、レビューテーブルが約50万件であった。メニューテーブルの件数とレビューテーブルの件数は、大きく違っていて、レビューを書かれていないメニューが多く含まれていることがわかった。地域度の算出に必要な話題性や満足度の算出には、レビューテーブルの中のカラムであるメニューの推薦者数や、メニューの評価点を使用する必要がある。そのため、地域度算出の対象となるメニューは、レビューテーブルが存在するメニューであることと考える。尚、地域度算出の際に、全体のメニュー数を扱う項目が存在するが、そのとき、ぐるなびから取得した全メニュー数を全体のメニュー数とみなすため、全体のメニュー数は約297万件として算出する。

(注2) : http://api.gnavi.co.jp/api/service.htm.

メニューテーブルは約 297 万件あり、その中には複雑な体系をなしているメニューもある。表 2 に複雑なメニューの体系を上げる。ぐるなびは店舗視点で記載されている為、ホームページが店舗の意向が表れており、メニューを目立たせようとするために「◆」や「■」などの記号や、「!」などのマークなど強調表現が多く見られる。その為、メニューの体系を分析する必要があり、予備実験としてメニューの完全一致で出現回数を分析する。

表 2 メニューの体系一例

メニュー id	メニュー
197535	■鰻ひつまぶし
1979644	◆とりひつまぶし
2049980	■秋鮭とイクラの親子ひつまぶし風

そこで、メニューを出現回数が多い順に上位 50 件を表にしたものを表 3 に示し、出現回数の多い上位 50 位までのメニュー分布を、図 3 に示す。ソフトドリンクやカクテルなど飲料系の総称や、ウーロン茶やコーラなど飲料系のありふれたメニューが見られた。また、枝豆やシーザーサラダなど複数の店舗で被りがありそうな定番メニューが多く見られた。これは飲料系メニューや定番メニューはありふれたメニューであるだけでなく、頻繁にチェーン店舗などに存在しているメニューであることが要因だと考えられる。

メニューの出現回数の全体の分布を、図 4 に示す。全体では約 297 万件のメニューを取得したが、出現回数で数えるとメニューの種類は約 100 万件であることがわかる。全体で見ると、メニューの出現回数は 1 位のソフトドリンクから上位約 20 万件までが出現回数 2 回以上であることがわかる。出現回数 20 万位以降のメニューは、出現回数 1 回のメニューが続く、全体の約 8 割のメニューが出現回数 1 回であることがわかる。このことからメニューの完全一致から出現回数を見た場合、差がつかないことがわかる。

逆にひつまぶし、ラーメンなどのよく見かけるありふれたメニューにおいて、メニューの出現回数を考えると、出現回数が多過ぎてしまい、ありふれたメニューでも、出現回数の差がつかないことが考えられる。よって、メニューを体系化する必要があるといえる。

### 3.2 メニューの体系化

地域度の評価式を算出するには、メニューの文字情報を体系化する必要がある。メニューの体系化によって、メニューの中でも頻出の特徴語を抽出する必要がある。手順として、図 5 のような形でメニューの体系化を行う。ありふれた一般的な名称であるひつまぶしのようなメニューを元に、メニュー DB に直接 SQL 文を投げることによって、メニューの中で頻出の特徴語を抽出する。SQL 文の検索文を表 4 に示す。メニューや、出現回数を問い合わせることで、手動でメニューの頻出の特徴語を探る。

メニューの体系化の手順を図メニューの体系化を行うにあたって、メニューにはメニューの特徴を表す単語があると考え、それらの単語をメニュー枝と呼ぶ。例えば、鰻ひつまぶしであ

表 3 出現回数の多いメニュー上位 50 件

順位	メニュー	件数	順位	メニュー	件数
1	ソフトドリンク	4496	26	ジンジャーエール	1676
2	枝豆	4243	27	おにぎり	1675
3	ウーロン茶	3270	28	角ハイボール	1674
4	キムチ	3185	29	サラダ	1608
5	チャンジャ	3014	30	ガーリックトースト	1607
6	生ビール	2928	31	お茶漬け	1603
7	カクテル	2727	32	グリーンサラダ	1576
8	シーザーサラダ	2624	33	アイスコーヒー	1535
9	オレンジジュース	2416	34	マルゲリータ	1524
10	冷奴	2360	35	ビール	1462
11	ポテトフライ	2325	36	たこわさび	1420
12	日本酒	2300	37	ペペロンチーノ	1404
13	ライス	2237	38	焼酎	1386
14	ワイン	2175	39	しいたけ	1386
15	フライドポテト	2158	40	ビーフカレー	1370
16	梅酒	2131	41	豚キムチ	1353
17	ハイボール	2075	42	手羽先	1352
18	サワー	2038	43	レバー	1347
19	カルボナーラ	2021	44	ポテトサラダ	1341
20	コーラ	2006	45	ごはん	1330
21	冷やしトマト	1735	46	シーフードカレー	1318
22	パニラアイス	1715	47	ハラミ	1315
23	デザート	1710	48	生ハムサラダ	1304
24	ウィスキー	1695	49	ししとう	1288
25	ウーロンハイ	1689	50	白菜キムチ	1287

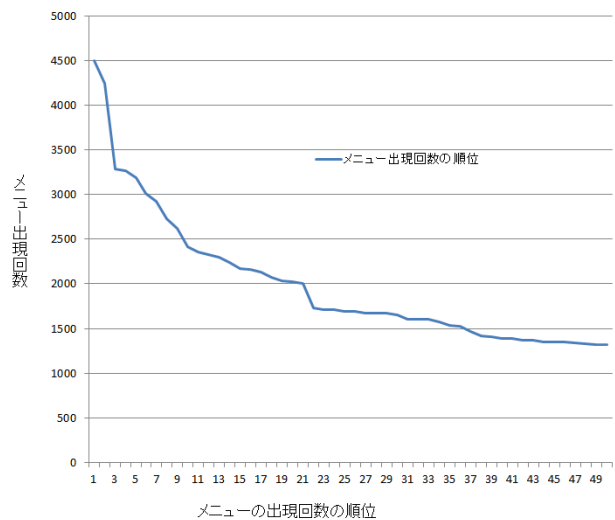


図 3 出現回数の多い上位 50 位までのメニュー分布

表 4 メニュー DB に入力する SQL 文

用途	SQL 文
メニューの単語抽出	select * from テーブル名 where カラム名 like '%メニュー%';
メニューの出現回数抽出	select カラム名,count(*) from テーブル名 group by カラム名 order by count(*) DESC;

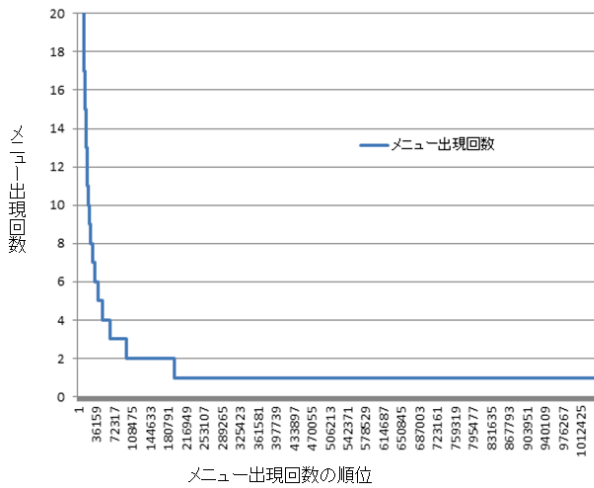


図 4 メニューの出現回数の全体分布

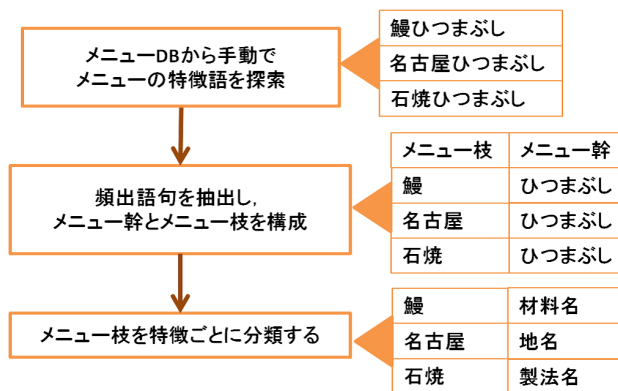


図 5 メニューの体系化の流れ

れば、鰻、ひつまぶしはそれぞれメニュー枝である。メニュー枝には以下のような特徴があると考えられ、地名、店舗名、材料名、製法名、その他名のようなものが挙げられる。

- メニューの地名であること。  
名古屋、浜松、浦和、浅草など
- メニューの店舗名であること。  
大戸屋、安楽亭など
- メニューの材料名となるものであること。  
鰻、鶏、穴子、鯛、海鮮など
- メニューの製法名であること。  
石焼、茶漬け、炙り、蒲焼、照り焼きなど
- メニューのその他名であること。  
うまい、特製、特盛など

次にそのメニューの核となるようなメニュー枝をメニュー幹と呼ぶ。メニュー枝が一つのみ存在する場合、そのメニュー枝がメニュー幹となる。メニュー枝が複数存在する場合、数あるメニュー枝の中で、そのメニューの特徴になる語をメニュー幹と考える。表 4 のやり方でメニューを抽出し、ひつまぶしがメニュー幹であるメニューを体系化したものを、図 6 に示すと、メニュー幹であるひつまぶしは 264 件ある。鰻と共起しているひつまぶしが 34 件、鰻と石焼と共起しているひつまぶしが 2

件、という見方になっている。「なし」はその左のメニューと共起する語がないメニューであり、例えばメニュー幹であるひつまぶしの右のメニューは、共起する語がなく、ひつまぶしというメニュー幹であることを表している。

メニュー名	件数	単語		単語	
ひつまぶし	264	なし	72	なし	72
		鰻	34	なし	24
				石焼	2
				名古屋	2
				炙り	1
				茶漬け	1
				浜松	1
				江戸前	1
		うなぎ	29	なし	10
				石焼	2
				蒲焼	2
				茶漬け	1
				特製	1
		海鮮	20	なし	17
				石焼	1
				弁天	1
		石焼	13	うなぎ	3
				海鮮	2
				鰻	2
				穴子	2
				なし	1
				鯛	1
				茶漬け	1
				照り焼き	1
				鶏	1
				浜松	1
		鶏	13	なし	7
				石焼	1

図 6 ひつまぶしのメニュー体系

ひつまぶしは鰻の蒲焼を用いた日本の料理で、名古屋市周辺で多く食べられているメニューである。鰻の蒲焼を細かく刻んでご飯に混ぜて食べる料理であり、製法名としても知られる。ひつまぶしのメニュー枝を分析していくと、上記の 5 つの形で分類することができると考えられる。名古屋や浜松、浦和、浅草などは、ひつまぶしと共起している料理の地名の名称であるので、本論文では地名とする。備長等の店舗名は、店舗名とする。鰻や鶏、穴子、鯛などは、ひつまぶしと共起している料理の材料名であるので、材料名とする。石焼、茶漬け、炙り、蒲焼などは、ひつまぶしと共起している料理の製法名であるので、製法名とする。特製などはそのメニューを強調しているメニュー枝であるので、その他名とする。上記のメニューの体系化の流れから、地域度の評価式を定義し、算出する。

### 3.3 地域度の評価手法

上記までのメニューの体系化を踏まえ、メニュー幹と共起するメニュー枝とのセットをメニューとし、以下の地域度の評価式を定義する。評価式は以下に示す地域性・話題性・満足度の 3 要素から構成されるものとする。

#### 3.3.1 地域性

地域性は、以下の 3 項目から構成される。地域性はその土地独自である指標であり、メニュー・店舗・地域の 3 項目から出現頻度を見て評価する。

##### メニュー $m_i$ の出現頻度 $M_i$

全国におけるメニュー  $m_i$  の出現頻度を式 (1) に示す。  $m$  は総メニュー数、  $mf(m_i)$  はメニュー  $m_i$  の出現総数を示す。

$$M_i = \log \frac{m + 1}{mf(m_i)} \quad (1)$$



メニュー  $m_i$  のある店舗の出現頻度  $S_i$

メニュー  $m_i$  を提供する店舗の全国における出現頻度を式 (2) に示す.  $s$  は総店舗数,  $sf(m_i)$  はメニュー  $m_i$  の出現する店舗総数を示す.

$$S_i = \log \frac{s+1}{sf(m_i)} \quad (2)$$

メニュー  $m_i$  を提供する店舗が存在する地域の出現頻度  $L_i$

メニュー  $m_i$  を提供する店舗が存在する地域の全国における出現頻度を式 (3) に示す.  $l$  は全地域数,  $lf(m_i)$  はメニュー  $m_i$  の出現する地域総数を示す.

$$L_i = \log \frac{l+1}{lf(m_i)} \quad (3)$$

以上の式 (1)~式 (3) の総和を求めることで, 地域性を定義し, 以下のように評価式 (4) を示す.

メニュー  $m_i$  の地域性  $Mlocal_i$

$$Mlocal_i = M_i + S_i + L_i \quad (4)$$

### 3.3.2 話題性

話題性は, メニュー  $m_i$  の推薦者数の店舗内の全メニューの推薦者数との比率を見る.

メニュー  $m_i$  の推薦者数の店舗における全メニューの推薦者の割合  $Msuisen_i$

メニュー  $m_i$  がその店舗においてどれだけ関心を集めているのかを式 (5) に示す.  $cf(m_i)$  はメニュー  $m_i$  の推薦者数,  $csf(m_i)$  はメニュー  $m_i$  のある店舗のメニューの総推薦者数を示す.

$$Msuisen_i = \frac{cf(m_i)}{csf(m_i)} \quad (5)$$

### 3.3.3 満足度

満足度は, メニュー  $m_i$  の評価点の比率を見る.

メニュー  $m_i$  の満足度  $Mscore_i$

メニュー  $m_i$  が 5 点満点中の飲食の満足度を式 (6) に示す.  $Mscore_i$  は, メニュー  $m_i$  の評価点, 5 はメニューの評価点の最高点を示す.

$$Mscore_i = \frac{mscore(m_i)}{5} \quad (6)$$

### 3.3.4 地域度

以上の式 (1)~式 (6) から地域度の評価式  $L(m_i)$  を定義し, 以下のように評価式 (7) を示す.

メニュー  $m_i$  の地域度  $L(m_i)$

$$L(m_i) = pMlocal_i \times (1-p)(Msuisen_i + Mscore_i) \quad (7)$$

地域度  $L(m_i)$  は, 話題性  $Msuisen_i$  と満足度  $Mscore_i$  の和と地域性  $Mlocal_i$  を掛け合わせることで, 評価する. パラメータ  $p$  を設定し, 利用者がそれぞれの性質に重視するかで, それぞれの評価値の重みを変化できるようにする. 本論文では, 地域度の評価のために,  $p=0.50$  の場合の地域度を算出する.

## 4. 地域度の算出と考察

本章では, 3 章で説明した地域度の評価式 (7) を元に, パラメータの値ごとに地域度を算出する. 4.1 節で地域度を示し, 4.2 節で結果について考察を述べる.

### 4.1 各メニューの地域度の算出

3.3 節で提案した評価手法にて, メニュー幹やメニュー枝からなるメニュー  $m_i$  に対して, 地域度を算出する. ひつまぶし, ラーメンサラダ, チキンカレーというメニューに対し, 地域度を算出する.

#### 4.1.1 ひつまぶし

3.2 記載の通り, ひつまぶしは全部で 264 件あり, その中でメニューの出現回数を表すと図 6 になる. ひつまぶしと共に起しているメニュー枝は鰻が一番多くなっている. ひつまぶし自体が製法名になり替わることも多いため, 多様なメニュー枝と共に起していると考えられる. それぞれの組み合わせで地域度を算出した結果を, 表 5 に示す.

算出結果を見ると, 地名である名古屋, 浦和, 浅草が上位に來ていることや, 材料名が上位に多く來ていることがわかる. 特に名古屋に関しては地域度が他よりも高くなっている. これは, 名古屋という土地に局地的に存在していることに加えて, 他と比べてメニューの話題性を評価する  $Msuisen_i$  や満足度を評価する  $Mscore_i$  の値が高いことが影響している. 浦和や浅草のひつまぶしは, 名古屋のひつまぶしの地域度との間に大きな差がある. ひつまぶしが名古屋市周辺のメニューであることが影響していることが理由だと思われる. 浦和と浅草は, その土地にメニューが局所的に存在しているものの, 地名ではないので話題性や満足度を測るための利用者の評価数が少なくなっている. 地域性の値は高くなっても, 話題性と満足度の値が低くなってしまうために, 浦和と浅草のひつまぶしの地域度は名古屋のひつまぶしには及ばない結果となったものと考えられる.

また材料名である天ぷら, コーチン, 鯛, 鶏, 鰻, 海鮮が上位 10 件に多い. これはひつまぶし自体が調理のしやすさや, 他の料理と合わせやすいメニューであることも大きいと考えられる. 実際に多くの薬味やメニューと合わせて食べられているので, その概算が大きい. これらは存在している地域数も少ないため, 地域度が高くなっていると考えられる. しかし, 名古屋のひつまぶしとは違い, 話題性と満足度の値が良くなかったために, 名古屋のひつまぶしよりも低い地域度を算出していると考えられる.

#### 4.1.2 ラーメンサラダ

ラーメンサラダは主に北海道で食されるラーメンとサラダが組み合わさった食べ物である. 冷やし中華より具材の野菜の量が多いので, ラーメンというよりもサラダとして食されているが, ラーメンサラダをメニュー幹として, ラーメンサラダの地域度を算出する. ラーメンサラダは, メニュー DB の中に表 4 のように SQL 文を投げると, 256 件ある. ラーメンサラダは, そのままの名前であるラーメンサラダが最も多いが, メニュー枝の種類としては材料名である温玉や豚しゃぶ, 味噌, 地名である北海道や札幌, 博多, 台湾などが存在する. ラーメンサラ

表 5 ひつまぶしの地域度上位 10 件

順位	メニュー	地域度	メニュー枝
1	名古屋ひつまぶし	9.80	地名
2	天ぷらひつまぶし	7.75	材料名
3	浦和ひつまぶし	7.46	地名
4	コーチンひつまぶし	7.05	材料名
5	鯛ひつまぶし	6.88	材料名
6	浅草ひつまぶし	6.71	地名
7	ひつまぶし茶漬け	6.46	製法名
8	鶏ひつまぶし	6.20	材料名
9	鰻ひつまぶし	5.84	材料名
10	海鮮ひつまぶし	5.29	材料名

ダの地域度の算出結果のうち、上位 10 件を表 7 に示す。

算出結果メニュー枝別に見ると、地名は博多、台湾、札幌、北海道、材料名は鶏、味噌、温玉、海鮮、大根、製法名は和風であるラーメンサラダの地域度が、上位 10 件である。メニュー枝が地名であるメニューを 2 つに分けると、博多や台湾のラーメンサラダが高く、札幌や北海道の地域度は低くなっている。ラーメンサラダは主に北海道で食される食べ物であり、ラーメンサラダに北海道や札幌など、北海道を表す名称を、北海道以外の地域でも使用している。そのため、局所性が少なくなり、地域性が低くなったと考えられる。それに対して博多や台湾のラーメンサラダは、その土地にメニューの局所性があり、更に話題性や満足度の値も高いため、北海道や札幌のラーメンサラダよりも地域度が高くなっていると考えられる。

表 6 ラーメンサラダの地域度上位 10 件

順位	メニュー	地域度	メニュー枝
1	鶏ラーメンサラダ	8.08	材料名
2	台湾ラーメンサラダ	7.93	地名
3	博多ラーメンサラダ	7.92	地名
4	味噌ラーメンサラダ	7.56	材料名
5	温玉ラーメンサラダ	6.89	地名
6	札幌ラーメンサラダ	6.58	地名
7	大根ラーメンサラダ	6.57	材料名
8	海鮮ラーメンサラダ	6.49	材料名
9	和風ラーメンサラダ	6.32	製法名
10	北海道ラーメンサラダ	5.33	地名

#### 4.1.3 カレー

カレーは、インドとその周辺国で作られていた料理をもとに発展し、複数の香辛料を使って野菜や肉などのさまざまな食材を味付けした料理である。カレーは、メニュー DB の中に表 4 のように SQL 文を投げると、7702 件ある。カレーのメニュー枝の種類としては材料名であるチキン、地名であるインド、カシミール、製法であるキーマなどが存在する。カレーの地域度の算出結果のうち、上位 10 件を表 7 に示す。

算出結果を見ると、4.1.1 や 4.1.2 とは違って、チキンキーマカレーのような、チキンという材料名とキーマという製法名などメニュー枝が複数存在し、多岐にわたっているメニューが上位に来ていることが分かる。メニュー幹の他にメニュー枝が 1 つであるメニューは、上位 10 件には 2 件のみであるのに対

し、メニュー幹の他にメニュー枝が複数あるメニューは、上位 8 件である。メニュー枝が増えることによって、メニュー、店舗、地域の局所性が高まり、地域性が上がったことが影響しているものと考えられる。ランク外のメニューである、チキンカレーやグリーンカレー、キーマカレー、牛カレーが存在するが、それぞれの地域度は 3.63, 4.15, 4.28, 5.97 にとどまっている。しかしながら、チキンカレーにエッグという材料名のメニュー枝が追加されると地域度は 8.39 であり、またキーマカレーにチキンという材料名のメニュー枝が追加されると地域度は 7.59 にまで高くなる。同じようにチキングリーンカレーも地域度が 7.18, 飛騨牛カレーや石垣牛カレー、神戸牛カレーも 10.07, 9.53, 8.62 であり高くなっている。このことからメニュー枝を付ければ付けるほど、メニューの地域性は高まり、その影響で地域度も高まるということが分かる。

上位 3 件は地名 + 牛カレーであるメニューが来ている。これは地域性も高いこともさることながら、飛騨牛や神戸牛は高級牛肉であり、話題性や満足度も高い値であったことが理由であると考えられる。このことからひつまぶしやラーメンサラダの地域度に関してもメニュー枝に地名を含むメニューが上位に来ていたが、カレーでも同様であることが分かる。

表 7 カレーの地域度上位 10 件

順位	メニュー	地域度	メニュー枝
1	飛騨牛カレー	10.07	地名 + 材料名
2	石垣牛カレー	9.53	地名 + 材料名
3	神戸牛カレー	8.62	地名 + 材料名
4	エッグチキンカレー	8.39	材料名 + 材料名
5	サーガマトンカレー	7.97	材料名 + 材料名
6	昆布だし牛すじカレー	7.83	材料名 + 材料名
7	チキンキーマカレー	7.59	材料名 + 製法名
8	カシミールカレー	7.21	地名
9	チキングリーンカレー	7.18	材料名 + 材料名
10	チキンティカマサラカレー	6.97	材料名 + その他名 + 材料名

#### 4.2 考 察

4.1 では、メニューごとに地域度の算出結果を算出したが、ひつまぶしとラーメンサラダ、カレーの地域度を一緒にしてまとめたグラフを図 7 に示す。この中でメニュー枝が製法名であるものは、ひつまぶし茶漬け、和風ラーメンサラダのみであり、主にメニューに地名と、材料名がメニュー枝であるメニューの地域度が高くなっていることがわかる。

上位の飛騨牛カレーや名古屋ひつまぶし、石垣牛カレー、神戸牛カレーは地域度と話題性や満足度が高い値を取っている。飛騨牛カレーや石垣牛カレー、神戸牛カレーはすべてが高いことに加え、メニュー枝からなる名前である飛騨牛や神戸牛が高級食材であるなど、メニューが上質であることから話題性や満足度の向上にもつながっていると考えられる。名古屋ひつまぶしはひつまぶしが名古屋生まれのメニューであり、名古屋周辺に局所的にメニューがあることや、評判の店舗が数多く存在することが挙げられる。上位 10 件までにメニュー枝が複数のメニューが 6 件と過半数あり、メニュー枝が多くなるほど地域性

このように地域度を構成する中でも、メニュー枝の性質や、メニュー枝の組み合わせ方によって、地域度は大きく変化することがわかった。

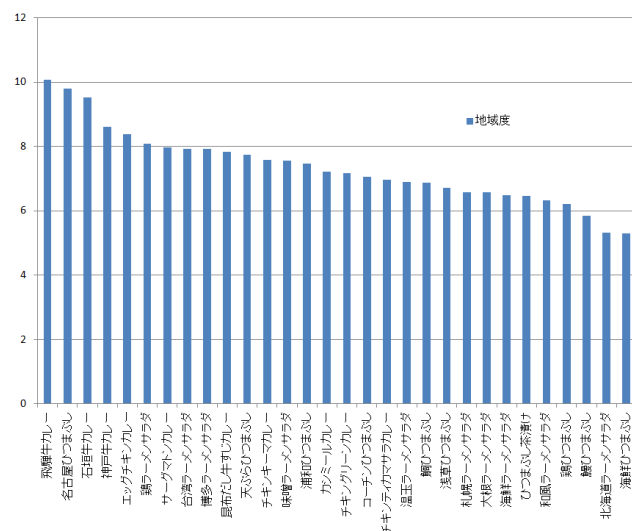


図 7 ひつまぶしとラーメンサラダとカレーの地域度

本論文の意義は、旅行や非日常的な出来事で外出をする際に、地域性・話題性・満足度の3つの視点からメニューを推薦することである。地域性は特定のエリアやメニュー等その土地独自である希少性を測るもので、空間的局所性を示していて、また話題性や満足度は、利用者がメニューの満足度や評価が高まりに感心しており、時間的局所性を示していると考えた。

地域度の評価式によって、ひつまぶし、ラーメンサラダ、カレーの地域度を算出すると、複数のメニュー枝を持つメニューの地域度が高くなり、中でも飛騨牛カレーや石垣牛カレー、神戸牛カレーのようなメニュー枝が地名であるものを含むメニューの地域度が高くなる傾向があることが分かった。またこれらのメニューは飛騨牛や神戸牛など高級食材が使われており、メニューが上質であることも影響して、話題性や満足度が高くなりやすいと考えられる。

これらの結果から、ひつまぶしやラーメンサラダ、カレーに対して、地域度の評価を行い、薦めるべきメニューが明らかにできたと考える。メニューの存在する地域や、メニューに対する推薦者数、評価点によって、地域度は大きく変化していた。今後は他の多くのメニューにも適用し、地域度がより確かなものであるかを検討していくべきと考える。

## 6. 謝 辭

本研究の一部は筑波大学図書館情報メディア系プロジェクト  
研究による助成を受けたものである。

## 文 献

- [1] 奥健太, 服部文夫, 地域限定性を考慮した情報推薦方式に関する基礎検討, Web とデータベースに関するフォーラム (WebDB Forum 2009), 1A-1
- [2] 西崎剛司, 奥健太, 服部文夫, 地域限定性を考慮した情報推薦における語句抽出の傾向分析とノイズの除去, 第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2011), B1-6
- [3] 手塚太郎, 近藤浩之, 田中克己, 混合ガウス分布を用いたウェブコンテンツの地域性推定とオブジェクトレベルローカルサーチ, 情報処理学会論文誌: データベース, Vol.1, No.1, pp. 13-25
- [4] 小林大祐, 井上潮, Web 上のレビュー情報からユーザが重要視する製品の特徴を抽出する手法の提案, 第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2009), C6-4
- [5] Evgeniy Gabrilovich and Shaul Markovitch, Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis, International Joint Conference on Artificial Intelligence, pp. 1606-1611, 2007