

## レビュアーの生活圏の特定に基づく CGM 評判情報評価

小林 卓弥<sup>†</sup> 大島 裕明<sup>†</sup> 小山 聡<sup>†</sup> 田中 克己<sup>†</sup>

<sup>†</sup> 京都大学大学院情報学研究科社会情報学専攻

〒 606-8501 京都市左京区吉田本町

E-mail: †{tkoba,ohshima,oyama,tanaka}@dl.kuis.kyoto-u.ac.jp

**あらまし** CGM における評判情報からレビュアーの「生活圏」を抽出し、それに基づくレビュー情報の評価方法を提案する。今回は飲食店評価をその対象情報とし、実際にレビュアーの「生活圏」の特定を行った。今後は「生活圏」内のレビューと「生活圏」外のレビューは異なる情報価値を有しているという考えの下に、複合的に「生活圏」を利用した、レビュー情報の評価方法を検討していく。

**キーワード** CGM, Blog, Web マイニング, 位置情報

## Evaluation of Reputations in CGM based on Identification of Reviewer's Activity Area

Takuya KOBAYASHI<sup>†</sup>, Hiroaki OHSHIMA<sup>†</sup>, Satoshi OYAMA<sup>†</sup>, and Katsumi TANAKA<sup>†</sup>

<sup>†</sup> Department of Social Informatics, Graduate School of Informatics, Kyoto University

Yoshida-Honmachi, Sakyo, Kyoto, 606-8501 Japan

E-mail: †{tkoba,ohshima,oyama,tanaka}@dl.kuis.kyoto-u.ac.jp

**Abstract** We propose a method for evaluating reviews based on identification of reviewer's activity area by using reputations in CGM. This time we use restaurant reputations in CGM, and identify a reviewer's activity area. We think the review inside reviewer's activity area has different information value from the review outside. In the future, we will discuss how to evaluate reviews by considering complex activity areas.

**Key words** CGM, Blog, Web Mining, Location Information

### 1. はじめに

近年、CGM(Consumer Generated Media)と呼ばれる消費者によるメディアが発展を遂げている。CGMとは、個人の情報発信をデータベース化、メディア化したWebサイトのことで、例として口コミサイト、ブログ、SNSなどが挙げられる。このメディア上で発信された情報には、消費者の実体験や生の声が存在し、それらレビュー情報を活用すべく様々な研究が行われており、例えば鈴木ら[1]による評判情報が肯定意見か否定意見かを判定するものなどが挙げられる。しかし、これら研究はいずれも自然言語処理により評判情報を抽出するのみにとどまっており、レビュー対象に対するレビュアーの背景を考慮したものは少ない。例えば、携帯電話に熟知した人が携帯電話に関するレビューを行った場合と、カメラに熟知した人が携帯電話に関するレビューを行った場合とは、異なる情報価値を有していると考えられる。

一般消費者が実際に口コミ情報などを参考にする際にしばしば行われる行動には、レビュアーが過去に行ったレビューの対象やレビューの内容から、レビュアーがどのような人物である

かを推測し、レビュー内容が信頼に足るものであるかを判定したり、自分と興味や関心が近いレビュアーを探すといったものがある。例えば過去に行われた携帯電話のレビューを参照し、好きな携帯電話のメーカーが一致するか調べるといったことである。この際、レビュアーが信頼できるかどうか調べるために各レビュアーのレビュー文章を読むという作業は、ユーザにとって大きな負担であり、レビュアーの有する背景を提示することでこれを軽減するシステムが望まれる。

そこで本論文では、レビュアーの通常の行動範囲や興味を有する対象範囲などの特定を行い、その情報を用いてレビュー情報の評価を行う手法を提案する。レビュアーの通常の行動範囲や興味を有する対象範囲などは「生活圏」と呼ぶ。「生活圏」というと、一般的には地理的な行動範囲ととらえられるが、我々は地理的な範囲だけでなく、興味・関心の範囲といった情報もこの「生活圏」として考える。

まず、レビュアー一人ひとりがこれまで行ってきたレビュー情報を集約することにより、レビュアーの地理空間や情報空間における「生活圏」を抽出する。それらをレビュー情報と共にユーザに提示することによって、ユーザがレビュアーを調査す

	基礎集落圏	一次生活圏	二次生活圏	地方生活圏
円域範囲	半径 1~2km	半径 4~6km	半径 6~10km	半径 20~30km
時間距離	老人・幼児の 徒歩限界 15~30 分	自転車 30 分 バス 15 分	バス 1 時間以内	バス 1~1.5 時間
中心都市及び 中心部人口	1 千人以上	5 千人以上	1 万人以上	15 万人以上
中心部の施設	児童保育, 老人福祉等の 福祉施設	役場, 診療所, 集会場, 小中学校等基礎的な 公共的施設	高度の買い物ができる商店街, 専門医をもつ病院, 高等学校等の地方生活圏, 中心都市の広域利用施設に準じた施設	総合病院, 各種学校, 中央市場等の 広域利用施設

表 1 標準的な圏域構成 [10]

る手間を省き、レビュー情報を評価する支援を行うことができる。さらに、ある商品や事柄に対して複数のレビュアーからレビューがなされているときには、レビュアー達の属性を集約した情報を提示することによって、どのようなレビュアー達からレビューされているのかといった傾向などを知ることができるようになる。このようにしてユーザが CGM 評判情報を評価する際の支援を行っていく。

本論文の以下では、まず 2 節で我々の考えるレビュアーの「生活圏」について述べ、3 節でその特定手法を示す。4 節で「生活圏」を考慮した評価について考察する。そして 5 節で関連研究を紹介し、最後に 6 節でまとめと今後の課題について述べる。

## 2. レビューアーの生活圏

本論文における、レビュアーの「生活圏」について考察する。「生活圏」とは何か、まず、様々な学問分野における定義から紹介する。これらは“生活圏”あるいは“生活空間”などとして論じられているものである。次に本論文における「生活圏」について述べ、最後に CGM 評判情報のデータモデルについて解説する。

### 2.1 生活圏

生態学の分野ではビオトープ (Biotop) と呼ばれる概念があり、“生活圏”や“生息空間”と訳される。それは生物個体あるいは個体群が存在できる環境条件を備える地域、あるいは生息場所と定義されるものである [9]。

地理学においては生活空間という語を、時間地理学的な生活活動の軌跡としての空間と定義し扱っている。そしてそれは空間的行動に制約がなく、個々人に帰属し、人間の数だけ存在している。

地学の分野では生存圏、あるいはレーベンスラウム (Lebensraum) と呼ばれる概念があり、それは国家が自給自足を行うために必要な政治的支配が及ぶ領土のことである [10]。ドイツでは、この語は各個人のテリトリーとして捉えられる。

社会学においては公共圏と呼ばれる概念があり、それは人間の生活の中で他人や社会と相互に関わりあいを持つ時間や空間、または制度的な空間と私的な空間の間に介在する領域のことである。

建築学・都市計画における“生活圏”は第一に、人々の住居を中心とした近隣生活の行動範囲である。近隣生活とは、人々

が互いに顔見知りであるということから成り立つ相互扶助的な地域生活である。具体的には基礎集落圏、一次生活圏、二次生活圏、地方生活圏といった語で定義される。その具体的な標準数値を表 1 に示す。第二に、都市生活圏域であり、労働や教育、買い物、レクリエーションなど、人々の日常生活が一日のうちで完結するような生活範囲を示す。これらはいずれも、交通体系、そしてその所要時間の影響を受ける。そのため、例えば交通体系が発達している地域とそうでない地域とでは、地理的な範囲に差が生まれる。

ここで、一般的に浸透している「生活圏」という語について考える。現在、広域生活圏、防災生活圏、地方生活圏といったように行政用語として「生活圏」という語が使用されるが、第二次世界大戦後の日本において「生活圏」という表現が一般化したのは、1969 年に経済企画庁が発表した「新全国総合開発計画」によるところが大きい [10]。そのため、一般的にイメージされる「生活圏」の定義は建築学・都市計画のそれが近いであろう。

### 2.2 本論文における生活圏

以上を踏まえ、本論文では、「生活圏」を“日常生活における活動範囲”と考える。これは、一般的にイメージされる、建築学・都市計画における「生活圏」を包含するものであるが、さらに、生活を構成するものとして、食事、労働、余暇、居住、服飾なども考慮する。一般的に「生活圏」というと地理的「生活圏」がイメージされるが、例えば好きな食べ物といった好物的「生活圏」、余暇に行く趣味といった趣味的「生活圏」なども定義できる。このように本論文では「生活圏」を多面的なものとする。

### 2.3 CGM 評判情報データモデル

ここで、本論文で対象とする、CGM 評判情報のデータモデルについて解説する。そのデータモデルを図 1 に示す。

各レビュアーは、レビュー対象の商品についてレビューを行う。このレビュー対象の商品には、家電、CD、書籍、映画、飲食店など様々なものがあり、それらは名称やカテゴリ、価格といった属性を持つ。レビュアーは性別、年齢、居住地、職業といった属性を持ち、レビューが行われると、投稿日、評価値、文章、使用金額といった評判情報データが作成される。

### 2.4 CGM 評判情報から見たレビュアーの生活圏

レビュアーの居住地属性を見ると、一般的にイメージされる地理的「生活圏」を求めることができる。例えば居住地が“京

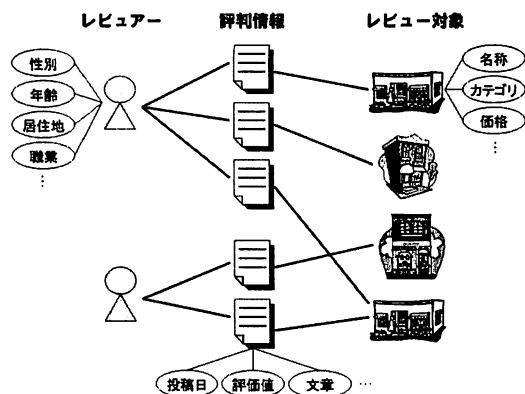


図 1 CGM 評判情報データモデル

都府京都市左京区”であることが分かれば、地理的「生活圏」をイメージすることができるだろう。

だが、我々はレビュー対象の情報を利用し、レビューアの「生活圏」を扱っていく。例えば、ある人の音楽的「生活圏」はビジュアル系バンドであるとか、ある人の書籍的「生活圏」は推理小説であるとか、ある人のコスメ的「生活圏」は 2,000 円以下であるということが、レビュー対象の属性を集約することで取得できる。

このように、本論文ではレビューが行われたレビュー対象の属性を集約することで、レビューアの「生活圏」を多面的に捉えていく。

### 3. 生活圏の特定

#### 3.1 対象とするデータ

本論文では、飲食店を対象とした CGM 評判情報を扱う。そのような飲食店レビューを含む CGM サービスとして、一般的なブログサービス [12]~[15]、食べログ [16]、ぐるなびみんなの口コミ [17]、hotpepper [18] などが挙げられる。本論文では、食べログを利用して実験を行う。この食べログの特徴として以下が挙げられる。

- 他サービスに比べ、レビュー数、レビューア数が多い
- レビューアによるレビュー対象店舗の追加が可能
- レビューアによる店舗評価値がある
- レビューアのプロファイルとして年代、性別、居住地がある

また、飲食店レビューアについては次のような「生活圏」が考えられる。

- 地理的な行動範囲
- 興味カテゴリの偏在性
- 飲食価格帯

#### 3.2 レビュー対象集約によるレビューア的生活圏

##### 3.2.1 地理的な行動範囲

まず、飲食店の位置情報を利用して、レビューアの地理的な行動範囲を推定する。レビュー対象の飲食店が持つ情報には、店名、住所、位置情報（緯度・経度）、電話番号、営業時間、

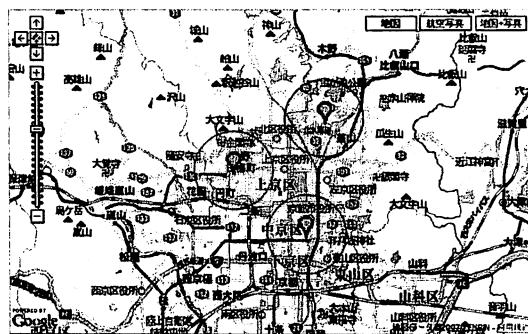


図 2 レビューア S の地理的行動範囲

カテゴリ、予算などがある。ここでは、レビューした飲食店の緯度・経度を取得し、地図上の距離を利用してクラスタリングを行う。手法は階層型クラスタリングとし、要素間の距離にはユークリッド距離を、クラスタ間の距離には重心法を用いる。そして閾値をその地域の時間距離を考慮した距離とし、一定以上の要素数を持つクラスタを「生活圏」として扱うものとし、そのクラスタが表す行動範囲の重心を求める。このアルゴリズムを以下に示す。

- (1) すべてのレビュー対象の緯度・経度を取得
- (2) それらを階層的クラスタリング（ユークリッド距離，重心法）
- (3) 距離の閾値でクラスタに分類
- (4) 一定以上の要素数を持つクラスタについて重心を求める
- (5) 閾値とした距離を半径，求めた重心を中心とする円で囲われる範囲を地理的な行動範囲とする

あるレビューアから取得した地理的行動範囲の例を図 2 に示す。今回は実験的に、重心からの行動範囲は自転車圏域と考え、閾値を 3km、またクラスタとして扱う要素数は 10 以上とした。このレビューアからは地理的な行動範囲の重心を三つ取得することができた。自転車圏となる 3km を点線の円で合わせて表す。

レビューアのプロファイルとして居住地を持つ場合も存在し、それから地理的「生活圏」のイメージを得ることはできるが、このように位置情報を利用することで、より詳細な「生活圏」が取得可能となる。また、通学先、通勤先、アルバイト先などが都道府県をまたぐ場合、そのレビューアの「生活圏」は居住地の都道府県を越えたものになると考えられる。そのため、このような位置情報の利用には有用性がある。

##### 3.2.2 興味カテゴリの偏在性

次に、レビューアの興味カテゴリの偏在性を特定する。これは、レビュー飲食店のカテゴリ分類を利用することで取得可能である。例えばレビュー 30 件中、和食 5 件、洋食 3 件、ラーメン 20 件、カフェ 2 件であれば、ここからそのレビューアはラーメン好きであることが予想される。レビューアによってレビュー件数が異なるため、ここでは各カテゴリのレビュー割合を求めるものとする。また、カテゴリによって店舗数が異なる

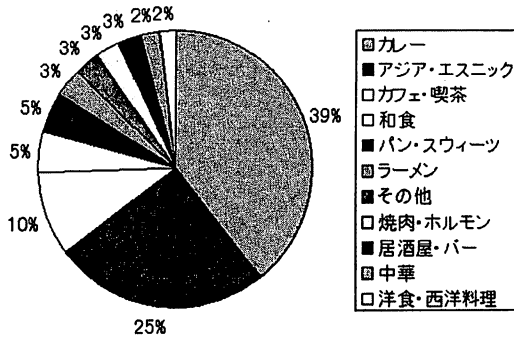


図3 レビュー P の興味カテゴリの偏在性

ため、レビューの出現確率が異なる。そこで、全国における各カテゴリの店舗数で正規化を行うものとする。このアルゴリズムを以下に示す。

- (1) すべてのレビュー対象のカテゴリを取得
- (2) カテゴリごとに件数を求める
- (3) レビュー者の総レビュー数、全国における各カテゴリの店舗数で正規化
- (4) その特徴ベクトルを興味カテゴリの偏在性とする

あるレビュー者について興味カテゴリの偏在性を特定した例を、図3に示す。件数で見ると和食カテゴリレビューが他のおよそ倍以上存在し、一見、和食に興味があるように思えたが、正規化により最も興味があるのはカレーであることが分かった。

### 3.2.3 飲食価格帯

最後に、レビュー者の飲食価格帯を特定する。レビュー情報の使用金額情報と、その利用時間帯(昼・夜)をレビュー者が作成した評判情報から取得する。“昼・0円~999円”などと、利用時間帯と使用価格帯の組で特徴ベクトルを作成し、その件数をカウントする。これも同様に、レビュー者によってレビュー件数が異なるため、使用金額の割合を求めるものとする。アルゴリズムを以下に示す。

- (1) 利用時間帯と使用金額を取得
- (2) 利用時間帯と使用金額の組で件数を求める
- (3) レビュー者の総レビュー数で正規化
- (4) その特徴ベクトルを飲食価格帯とする

## 4. 生活圏を考慮した評価

### 4.1 レビュー情報へのレビュー者の生活圏情報提示

我々は、ユーザがレビュー情報を参照している際、そのレビュー者の「生活圏」を提示することが必要であると考え、レビュー者の有する背景を提示されることによって、ユーザはレビュー情報に対して独自の評価を行うことができる。例えば、参照しているレビュー者 S のレビュー店舗の位置が図2の範囲に含まれない場合、その店舗はレビュー者 S にとって遠くても行く価値があると考えられるかもしれない。また、図3が示す興味カテゴリと、参照中のレビュー情報のカテゴリが一致した場合、そのレビュー情報の内容は信頼できると考えるかもしれない。

このように、まず各レビュー者の「生活圏」を提示すること

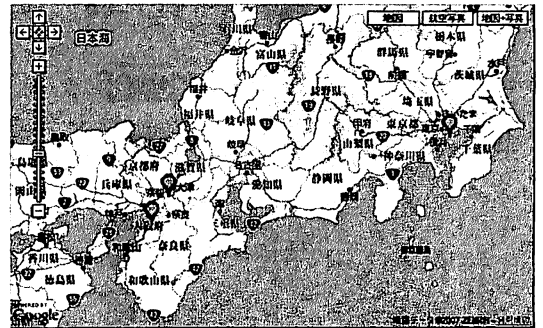


図4 店舗 A のレビュー者の居住地

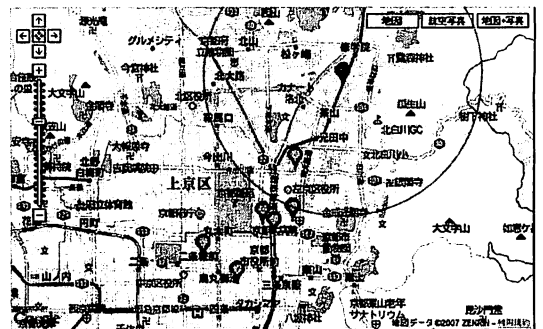


図5 店舗 A のレビュー者の京都「生活圏」

で、ユーザがレビュー情報を評価する際の支援を行うことができる。

### 4.2 レビュー対象へのレビュー者の集約情報提示

さらに、その各レビュー者の「生活圏」を集約して提示することを提案する。ここに、京都のある店舗 A のレビューを行ったすべてのレビュー者について、レビュー者の属性として取得できた居住地を集約したものを図4に示す。プロット点は都道府県庁所在地とした。このような図を提示することで、ユーザは全国の人が訪れている店舗かどうかを知ることができる。

さらに、そのレビュー者達の京都における地理的な行動範囲の重心を抽出した結果を図5に示す。3.2.1節と同様に、閾値を3km、クラスタとして扱う要素数は10以上とした。また、右上の赤いプロット点が店舗 A の位置であり、それを中心とした自転車圏3kmを点線の円で合わせて表す。レビュー12件中、居住地を取得できたのは9件で、内訳は京都府4件、大阪府2件、東京都3件であり、居住地が京都府外のレビュー者に関して京都府内に飲食「生活圏」が認められたものもあった。

これにより、普段どのような地域のレビューを行っている人がその店のレビューをしているかという情報を提示することができ、レビュー情報を探る際のユーザの支援が可能になると考えられる。例えばユーザは図5から、店舗 A は遠くから時間をかけてでも行きたい店であるとか、店舗 A は車で行く店であるといったことを推測するかもしれない。また、すべてのレビュー者が店舗 A を地理的「生活圏」に含んでいたら、店舗 A は近くに位置するから飲食しているだけかもしれないと見るかも

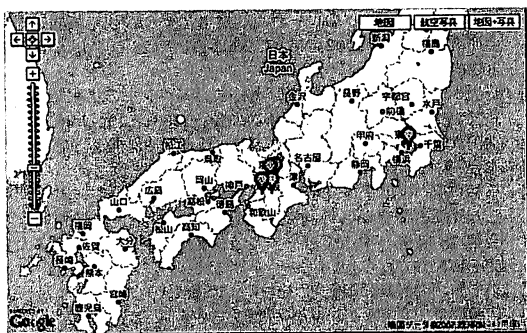


図 6 店舗 G のレビューアの居住地

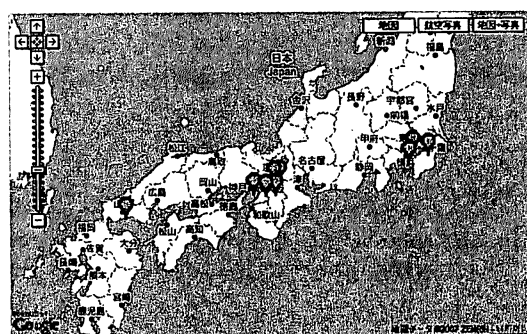


図 7 店舗 I のレビューアの居住地

しれない。

また、カテゴリが同一で、飲食店の位置、食ベログにおける総合評価値、レビュー件数が近い京都の飲食店 G と I について、レビューアの居住地で集約を行った例を図 6,7 に示す。プロット点は都道府県庁所在地とした。取得できた居住地件数は異なったが、店舗 G は近畿地方 29 件、関東地方 2 件であり、店舗 I は近畿地方 12 件、関東地方 10 件、中国地方 1 件であった。ユーザは、G は地元の人が行く店舗、I は全国的に有名な店舗と考えるかもしれない。

このように、レビューアの「生活圏」を集約することで、ユーザがレビュー対象を評価することが可能となる。

#### 4.3 考 察

今回の予備実験では、地理的な行動範囲を自転車圏とし、閾値を 3km とした。しかし、使用する交通体系によって各レビューアの行動範囲は、徒歩圏であったり、自動車圏であったりするであろう。そのためこの閾値は、レビューアの交通体系を推測して、ある程度動的に変化することが望ましい。また、図 2 のような地図だけではなく、その周囲の地名を表示することも考えられるため、そのような地名間の距離を閾値とすることも今後考えていく。同様に要素数に関しても、動的に細分化したクラスタ作成ができるよう考慮していく。

次に、情報空間における「生活圏」について考察する。地域によって、人口あたりの店舗数が少ないカテゴリの飲食店が存在し、その格差は最大 20 倍以上になる [11]。そのため、今回は全国単位でのカテゴリの店舗数で正規化を行ったが、より詳細

な地方ごとのカテゴリ店舗数で正規化を行うことが望ましい。

#### 4.4 生活圏の多面性を利用した評価

今後行っていく、「生活圏」を利用した評価方法について検討する。一般的に、「生活圏」内のレビュー情報価値は高い。しかし、地理的に見ると「生活圏」外だが、興味的に見ると「生活圏」内である店舗情報の評価をどう行うか。ある地理空間において、多数を占める店舗カテゴリが存在する場合、そのカテゴリに興味があるレビューアとないレビューアでは、そのレビュー情報の価値が異なると考えられる。例えば、中華料理の多い地域が「生活圏」であり、中華料理に興味を持っているレビューアと、日本料理に興味を持っているレビューアとは、地理空間における満足度が異なるであろう。その地理的「生活圏」において日本料理が少ない場合、後者のレビューアは、自らの行動範囲を離れ、日本料理の店舗を探し歩いているということが考えられる。我々はこのように、特定した「生活圏」を複合的に利用した評価手法を今後検討していく。

### 5. 関連研究

中島ら [2] はブックマーク作業における Web 閲覧プロセス (コンテキスト) に着目し、コンテキストにおけるユーザの背景をブックマークページの評価値として利用するシステムを提案している。これは、ブックマークするまでのコンテキスト内において、ブックマークした Web ページとの類似性が高いページを多く閲覧していればそのブックマーク行為に価値があるとしたものである。本論文の情報空間におけるユーザの「生活圏」に類するものであるが、我々はさらにレビュー対象での集約を行う。また、竹原ら [3] はニュース記事に対する意見を Blog から抽出し、Blog の書き手の興味を考慮して意見を提示するシステムの提案を行っている。本研究と同様に、書き手の過去の記事を利用してその書き手の興味を推測している。本研究ではこれに加え、地理空間における「生活圏」についても特定し、さらに多面的な集約を行う。

地理空間における「生活圏」と同様の研究として、山本ら [4] による、位置情報を持つ投稿情報に信頼度を付与するものが挙げられる。これはレビューアが情報を投稿した際、その周囲にも投稿情報があれば、その情報内容の信頼度は高いとしたものである。我々はさらに、レビューアの「生活圏」の多面性にも注目し、評価を行う。

馬ら [5] は、ウェブページがどの程度地域に密着しているかを判定するために、ウェブページ内容の偏在性、話題の偏在性とユーザの偏在性を考慮した、Web コンテンツのローカル度の定義を行っている。これは、ページに含まれる地域国名、組織名、地域一般名といった地理用語を利用したものである。しかし、詳細な地域名が無い、あるいは一般的でない地域においては、レビュー情報に期待する地理用語が含まれない。本論文の地理空間における「生活圏」は、レビュー情報の位置情報を利用しているので、そのような地域においても「生活圏」を抽出し、その詳細性を判断することが可能である。また、小林ら [8] は、ブログ記事における表層的特徴ならびに格フレームを利用し、ブログ文書から訪問場所名詞を抽出する手法を提案してい

る。メタデータとして位置情報を持たないブログ文書から「生活圏」を抽出する際には、このような手法により訪問場所を抽出することが必要になってくるであろう。

野呂ら [6] は、プログラマーの行動分析として、ブログ記事におけるイベントの生起時間判定を行っている。これはブログ記事からイベント文を抽出し、朝・昼・夕・夜・不明の 5 値で行動文の時間帯を判定するものである。例えば昼は勤務先周辺、夜は自宅周辺といったように、時間帯によって人の「生活圏」が異なる場合が考えられる。このようなイベント生起時間判定の研究は、ブログから地理的な行動範囲を抽出する際のメタデータとして重要になってくるであろう。また、池田ら [7] は一人称代名詞や機能語などを利用してブログの著者の性別推定を行っている。今回定義した「生活圏」というレビューの有する背景を提示する際、合わせてこのようなレビューのプロファイルを提示することが、ユーザのレビュー情報探索コスト軽減に繋がると考えられる。

## 6. まとめと今後の課題

本論文では、ユーザがレビュー情報を探索する際、各レビューの有する背景を提示することがユーザ負担を軽減すると考え、レビューの通常の行動範囲や興味を有する対象範囲などを「生活圏」と定義し、その特定を行った。

今後は、動的な閾値決定、要素数決定など、まず「生活圏」を特定するアルゴリズムの修正を行う必要がある。その上で、考察した「生活圏」を複合的に利用したレビュー情報の評価を行っていく。また、今回三つの「生活圏」を考えましたが、これらを統一的に見せるインタフェースについても考察していく。

## 謝辞

本研究は一部、文部科学省研究委託事業「異メディア・アーカイブの横断的検索・統合ソフトウェア開発」(研究代表者: 田中克己)、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」における計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」(研究代表者: 田中克己, A01-00-02, 課題番号 18049041)、および、「情報爆発に対応する新 IT 基盤研究支援プラットフォームの構築」(研究代表者: 安達淳, Y00-01, 課題番号: 18049073) 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] 鈴木泰裕, 高村大也, 奥村学, “Semi-Supervised な学習手法による評価表現分類”, 言語処理学会 第 11 回年次大会, 2005.
- [2] 中島伸介, 黒田慎介, 田中克己, “閲覧履歴を反映したコンテキスト依存型 Web ブックマーク”, 情報処理学会論文誌: データベース, Vol.43, No.SIG 5(TOD14), pp.23-36, 2002 年 6 月.
- [3] 竹原 幹人, 大島 裕明, 田中 克己, “Blog における書き手の興味を考慮した意見情報の提示手法”, 情報処理学会研究報告, Vol.2005, No.67, 2005-DBS-137, 2005 年 7 月.
- [4] 山本浩司, 片上大輔, 新田克己, 相場亮, 桑田仁, “地図上の投稿情報の信頼度”, 2006 年度人工知能学会全国大会 (第 20 回) 論文集, pp.3G1-3, 2006.
- [5] 馬強, 松本知弥子, 田中克己, “ページ内容と位置情報に基づ

く Web コンテンツのローカル度検出とその応用”, 情報処理学会研究報告, Vol.2002, No.67, 2002-DBS-128-69, pp.515-522, 2002 年 7 月.

- [6] 野呂太一, 乾孝司, 高村大也, 奥村学, “イベントの生起時間帯判定”, 言語処理学会 第 12 回年次大会, pp.837-840, 2006.
- [7] 池田大介, 南野朋之, 奥村学, “blog 著者の性別推定”, 言語処理学会 第 12 回年次大会, 2006.
- [8] 小林卓弥, 吉川正俊, “Blog の表層的特徴と格フレームを利用した Blog 文書からの訪問場所名詞の獲得”, 電子情報通信学会第 18 回データ工学ワークショップ第 5 回 DBSJ 年次大会 (DEWS2007), C3-2
- [9] 小学館『大辞泉』編集部, 松村明, “大辞泉”, 小学館.
- [10] 伊藤喜栄, “図説 日本の生活圏”, 古今書院.
- [11] “「牛丼」店舗密度は東高西低”, 日本経済新聞夕刊 (なるほどビジネス Map), 2006 年 10 月 25 日付.
- [12] livedoor ブログ,  
<http://blog.livedoor.com/>
- [13] FC2 ブログ,  
<http://blog.fc2.com/>
- [14] Yahoo! ブログ,  
<http://blogs.yahoo.co.jp/>
- [15] Ameba(アメブロ),  
<http://www.ameba.jp/>
- [16] 全国レストランガイド [食ペログ.com],  
<http://r.tabelog.com/>
- [17] ぐるなびみんなの口コミ,  
<http://kuchikomi.gnavi.co.jp/>
- [18] ホットペッパー.jp,  
<http://www.hotpepper.jp/>