

地域ブログサイト全量解析による Local Expert ユーザのランキング手法

三浦 惇貴^{†,a} 廣田 雅春^{†,b} 野澤 浩樹^{‡,c} 横山 昌平^{†,d}

† 静岡大学大学院情報学研究所 ‡ 静岡大学創造科学技術大学院/日本学術振興会特別研究員 DC

‡‡ 株式会社シーポイント

a) gs14043@s.inf.shizuoka.ac.jp b) dgs11538@s.inf.shizuoka.ac.jp

c) hiro@c-point.co.jp d) yokoyama@inf.shizuoka.ac.jp

概要 近年、日々の様々な活動にゲーム性を加えるゲーミフィケーションに注目が集まっている。本研究では、ブログサイトにゲーミフィケーションを導入するために、記事の言及している場所 (Venue) を発見し、各 Venue に詳しく言及しているブログユーザ (Local Expert) を、詳しくに基づいたランキング形式で提示する手法を提案する。その際、情報の鮮度を考慮するため、記事の投稿日時を考慮する。このようにゲーミフィケーションを導入することにより、ブログユーザにとっては、Venue についての意見や感想を含む記事を投稿するモチベーションとなり、Venue に関する情報を取得したいユーザにとっては、参考になる記事が増加することが期待される。我々は、浜松を中心とした地域ブログサイト [はまぞう] の記事全量を解析する事によって浜松市内の Venue を発見し、各 Venue の Local Expert をランキング形式で提示するシステムを構築した。

キーワード ユーザ評価, レビュー, ゲーミフィケーション, Venue, Foursquare

1 はじめに

近年、位置情報ゲームをはじめとする、Foursquare¹ などのゲーミフィケーションを取り入れた SNS に注目が集まっている。位置情報ゲームとは、携帯電話向けゲームアプリの一種であり、位置情報を利用したゲームのことである。ゲーミフィケーションは、SNS におけるブログユーザの投稿に対するモチベーションの向上や、ブログユーザ間のコミュニケーションの促進に効果が期待されている [1]。しかし、ゲーミフィケーションにより励起されるのは、既存のブログユーザの投稿に対するモチベーションであり、アーリーアダプターでないユーザの新規参加に対するモチベーションではない。つまり、既存のサービスにゲーミフィケーションをコールドスタートで導入した場合、アーリーアダプターを超えて、一般的なユーザに参加を訴求する事は困難を伴う。そのため、既存のサービスにゲーミフィケーションを低リスクで導入するには、システムをホットスタートさせる事が重要となる。ここでのリスクとは、ゲーミフィケーションを導入した新しいサービスが、すべてのブログユーザに継続して利用されるかということを用いる。また、既存のサービスにゲーミフィケーションを新たに導入しても、ブログユーザに負担が必要なゲームでは、ブログユーザは利用しない可能性がある。そこで、本研究では、SNS に馴染みのないユーザを含めた全ブログユーザが、無自覚のまま位置情報ゲームに参加している状態を作り上げ

ることにより、ゲーミフィケーションをホットスタートさせる手法を提案する。

本研究の目指すゲーミフィケーションは、Foursquare のように訪れた回数のみを利用する位置情報ゲームに加えて、記事に言及されている意見や感想も利用する。具体的には、場所 (店名やホテルなど) について言及した記事数と言及内容に基づいて、その場所について詳しくの度合いを算出し、ブログユーザをランキング形式で提示を行う。このようなゲーミフィケーションにより、意見や感想の記事を書くためのモチベーションになると考えられる。また、意見や感想の記事が増えることで、その場所について情報を取得したいユーザの参考になると考えられる。

本研究では、ブログサイトにおいて、位置情報ゲームを実現するために記事の言及している場所 (Venue) を発見する。発見された Venue について言及しているブログユーザの記事から感想を抽出し、Venue に対して詳しく言及しているブログユーザ (Local Expert) を選出することで、Venue に対する詳しくに基づいたランキングを作成する。また、Venue について鮮度の高い情報を優先するため、ランキング作成後に投稿された記事の重みを重視しつつランキングを更新する。これにより、ブログユーザは、記事を投稿するだけで、無自覚のままゲームに参加していることとなる。また、過去の記事を用いてランキングを作成することにより、システムをホットスタートさせることができる。本研究では、浜松を中心とした地域ブログサイト [はまぞう]² の記事の全量データ

Copyright is held by the author(s).

The article has been published without reviewing.

¹<https://foursquare.com/>

²<http://www.hamazo.tv/>

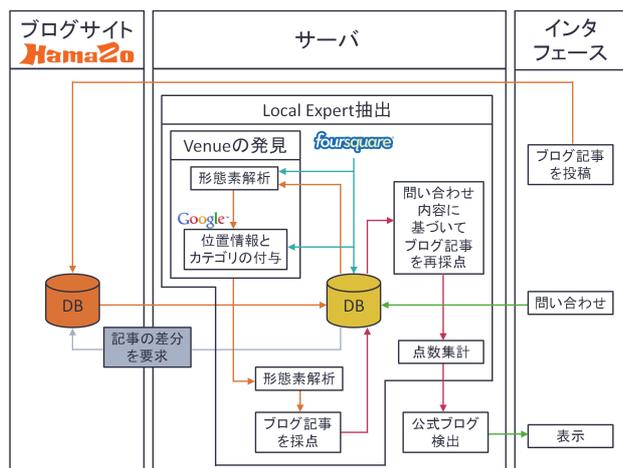


図1 システム構成図

を用いた実行例を示す。

2 関連研究

2.1 Venueの発見に関する研究

河野らの研究 [2] は, Venue を発見する際に, 形態素解析とキーフレーズ解析を行い, 地名を含んだ単語や, キーフレーズを Venue としている. また, 倉島らの研究 [3] は, 形態素解析を用いて, 地名・ランドマークと格助詞の共起に着目し, 格助詞と共起している単語列を Venue としている. さらに, 中嶋らの研究 [4] は, 場所を表す格助詞が出現する文を係り受け解析し, その格助詞の直前の名詞を Venue としている. これらの研究は, 形態素解析や, 係り受け解析に基づいており, 特定の品詞で表現された Venue 以外は発見できないという課題がある. そのため, 本研究では, 実際の地名や, 店名などの Venue を表す単語がユーザによって登録されている Foursquare を利用することで, その課題に対応する.

2.2 レビュー分析に関する研究

平山らの研究 [5] は, 商品の評価項目を自動で抽出する手法を提案している. 平山らの手法では, レビュー文に対して, 係り受け解析を適用し, 特定の品詞の語を評価項目として抽出している. また, 小西らの研究 [6] は, トピックモデルを用いて, 評価項目を推定する手法を提案している. 本研究では, 記事を扱うため, レビュー以外の文も数多く現れるため, これらの手法を適用した場合に, 適切な評価項目が得られないと考えられる. よって, 現段階では, 本研究では, 評価項目を恣意的に与えることで対応する.

2.3 ユーザランキングに関する研究

Baoらの研究 [7] は, Foursquare を用いて, Local Expert を抽出する手法を提案している. Baoらの手法は, Foursquare のチェックイン履歴とチェックイン回数を用いて, カテゴリ毎に Local Expert を抽出している. 本研

表1 形態素が1つの名詞でない Venue の例

Venue	形態素解析結果
にこまん馬	にこま (動詞)/ん (助動詞)/馬 (名詞)
ピッツァ・なお	ピッツァ(名詞)/・(記号)/なお (接続詞)
とんきい	とん (動詞)/きい(名詞)
中華の鉄人	中華 (名詞)/の (助詞)/鉄人 (名詞)

究では, Venue に対する詳しくさを考慮するため, チェックイン回数に加えて, Venue に対する詳しくさを考慮した Local Expert 抽出を行う.

稲垣らの研究 [8] は, 体験熟知度に基づいてブロガーのランキングを作成する手法を提案している. “マニア”などの単語の直前の名詞句を記事から抽出し, 名詞句の共起語と名詞句を含む記事の回数を基にブロガーのランキングを作成している. 本研究では, Venue に対して詳しく言及しているプロゲユーザーを抽出するため, “マニア”などの単語との共起を利用して Venue を発見することは困難である. よって, 前述したように, Foursquare から収集した Venue を形態素解析辞書に事前に登録することで対応する.

3 提案システム

本研究で提案するシステムの構成図を図1に示す. はじめに, Venue を発見するため, 記事を形態素解析し, その結果から Venue 候補を発見する. そして, Venue 候補に対して, 位置情報とカテゴリを付与する. その後, Venue のカテゴリに基づいて, 記事の採点を行う. 問い合わせ内容に基づき, 記事の時間を考慮し再採点を行い, Local Expert 抽出を行う. 最後に, ランキングに含まれる Local Expert から, 公式ブログを検出し, その結果をランキング形式でユーザに提示する. また, ランキング作成後は, 新たに投稿された記事の重みを高くしつつランキングを更新する.

3.1 Venueの発見

本研究で発見する Venue とは, 場所を表す単語のことを示し, 例として, 地域名である“浜松”や, 店名である“にこまん馬”などが挙げられる. 発見した Venue には, カテゴリを付与し, Local Expert 抽出を行う際の, Venue の評価項目の決定に利用する.

3.1.1 Venue候補の発見

記事に対して形態素解析を適用し, Venue 候補を発見する. 形態素解析には, Kuromoji³ を用い, 形態素解析辞書として, NAIST Japanese Dictionary⁴ を用いる.

Venue を含む記事に形態素解析を適用した場合, 表1に示すように, Venue が1つの形態素と判定されない

³<http://www.atilika.org/>

⁴<http://sourceforge.jp/projects/naist-jdic/>

場合がある。これに対処するために、Foursquare に登録されている Venue に含まれている name(F-Venue) を用いる。浜松市役所から半径 16km 以内の F-Venue を、Foursquare API⁵ を用いて収集し、形態素解析辞書に事前に登録する。また、F-Venue の登録は、以下のルールに従って行う。全角スペース、および半角スペースが含まれている F-Venue(e.g. “CLOCHE (クロッシュ)”) は、全角スペース、または半角スペースで区切り、別々に単語としての登録を行う。4 文字以下の英数字のみの単語や、1 文字のみの単語は登録を行わない。さらに、品詞情報が「地域」である単語や、単語の末尾が“店”(e.g. 浜松店) や“処”(e.g. そば処) である単語も登録を行わない(e.g. “ホテルコンコルド 浜松”は、“ホテルコンコルド”のみ、また、“そば処 津がわ”は、“津がわ”のみを登録)。“.”を含む単語は、“.”を含む単語と除去した単語の両方を登録する(e.g. “ピッツァ・なお”に対して、“ピッツァ・なお”と“ピッツァなお”両方を登録)。丸括弧を含む単語は、丸括弧内に記載されている単語を登録する(e.g. “CLOCHE (クロッシュ)”)に対して、“CLOCHE”と“クロッシュ”両方を登録)。以上のルールに従い登録した辞書を用いて形態素解析し、得られた名詞、または登録された単語(F-名詞)を Venue 候補とする。

3.1.2 Venue の決定

3.1.1 節で発見した Venue 候補に対して、次に述べる手順で位置情報、およびカテゴリ(e.g. “Japanese Restaurant”, “Sake Bar”)の付与を行う。

Venue 候補の品詞情報が「名詞」かつ「地域」の場合には、“楕形”のように地名が店名などの名前になっていることがあるため、Venue 候補が F-Venue に含まれている場合は、F-Venue の位置情報、およびカテゴリを付与する(F-Venue に用いた Foursquare の Venue の location、および categories を利用)。含まれない場合は、Google Geocoding⁶ を用いてジオコーディングを行い、Venue 候補に対して位置情報と“Neighborhood”というカテゴリを付与する。ジオコーディングの結果として、複数の位置情報が得られる場合がある(e.g. 単語“浜松”をジオコーディングした場合、東京都浜松町と静岡県浜松市)。そのため、本研究では、ジオコーディングの際に浜松市役所から半径 16km 以内の結果のみを利用する。その他の「名詞」は、F-Venue と一致した場合のみ、位置情報、およびカテゴリを付与する。F-名詞の場合は、該当の F-Venue の位置情報、およびカテゴリを付与する。この処理によって、位置情報、およびカテゴリが取得できたものを Venue とする。

⁵<https://developer.foursquare.com/>

⁶<https://developers.google.com/maps/documentation/geocoding/>

表 2 レストラン系の Venue の評価項目および類義語

評価項目	類義語の一部
味	風味, 旨味, こく
見た目	見映え, 表面, 見かけ
サービス	待遇, 接客, 値下げ
雰囲気	ムード, 佇まい, 風情
コストパフォーマンス (コスパ), ボリューム	量, 分量, 大きさ
酒, ドリンク	アルコール, ワイン, 焼酎

表 3 ホテル系の Venue の評価項目および類義語

評価項目	類義語の一部
部屋	個室, 座敷, ルーム
設備, アメニティ	備え付け, 施設, 設置
立地	場所, 敷地, 用地
接客, サービス	レセプション, 応対, 接待
食事, 料理	夕飯, ダイニング, ディナー
料金	値段, 代金, 払い

3.2 Local Expert 抽出

本研究で抽出する Local Expert は、Venue に対して詳しく言及しているブログユーザのことを示す。Local Expert は、Venue に関する記事を投稿した回数(チェックインポイント)や、Venue に関する記事が含む Venue に対する評価(記事ポイント)に基づいて、ブログユーザの Venue に対する詳しくさを評価する。

以下の式を用いてブログユーザ u の Venue に対する、スコア $Score_u$ を求め、 $Score_u$ の降順にブログユーザをランク付けする。

$$t_{iu} = \begin{cases} 0 & (dif_i < 0) \\ \frac{month - dif_i}{month} & (otherwise) \end{cases} \quad (1)$$

$$Score_u = \sum_{i \in I_u} a_{iu} * t_{iu} + \sum_{i \in I_u} c_{iu} * t_{iu} \quad (2)$$

ここで a_{iu} は記事 i に対するブログユーザ u の記事ポイント、 c_{iu} は i に対する u のチェックインポイント、 t_{iu} はブログの投稿日時による重み、 $month$ はランキングの算出に用いる記事の投稿日時の期間を月で表した数値、 dif_i はランキングの算出に用いる記事の投稿日時が最も新しい日付と i の投稿日時との月数の差である。

記事ポイント a_{iu} とは、ブログユーザ u の記事 i に対して、形態素解析を適用し、6 つの評価項目の単語や、評価項目の類義語が出現した場合、その記事を投稿したブログユーザに与えるポイントである。評価項目や評価項目の類義語が出現した場合、記事ポイントとして 1 点

表4 ランキング結果 (2010.7-2013.6)

rank	author	a_{iu}	c_{iu}	$Score_u$
1	★ ospitare	17.88	13.50	31.38
2	niimi	12.58	9.44	22.02
3	kakesta	11.94	5.16	17.11
4	blancjour3	9.63	3.86	13.49
5	leaf453	4.63	2.33	6.97

表6 ランキング結果 (2013.1-2013.6)

rank	author	a_{iu}	c_{iu}	$Score_u$
1	officesola	2.50	1.66	4.15
2	niimi	1.99	1.99	3.99
3	thepartyparty	1.00	1.00	2.00
4	bomblove	0.83	0.83	1.66
5	neolife	0.83	0.83	1.66

表5 ランキング結果 (2010.7-2010.12)

rank	author	a_{iu}	c_{iu}	$Score_u$
1	★ ospitare	17.00	12.49	29.49
2	blancjour	4.16	2.49	6.66
3	kakesta	1.16	1.16	2.33
4	sel500	1.16	1.00	2.16
5	sorapapa	1.16	0.83	2.00

を与える。各評価項目に対して、記事ポイントは最大1点とするため、1記事に対する記事ポイントは最大6点となる。評価項目は、注目 Venue(Local Expert 抽出を行う Venue) のカテゴリに基づいて定める (e.g. カテゴリが “Japanese Restaurant” の場合は、「味」, 「雰囲気」などの項目)。チェックインポイント c_{iu} とは、ブログユーザ u の記事 i の記事ポイント a_{ui} が 1 以上の場合、 $c_{ui} = 1$ となる。

3.3 公式ブログの検出

本研究では、Venue に詳しいブログユーザを抽出することで、Venue を実際に利用したブログユーザの感想を取得することも目的である。そのため、公式ブログのように Venue 側から発信された情報と、Venue を実際に利用したブログユーザの情報を区別するために、あるブログユーザのブログにおいて、1つの Venue に対する記事数が、総記事数の 15% 以上を占めている場合に、そのブログユーザを公式ブログのブログユーザとする。

4 提案手法によるランキング結果

3.1 節で発見した Venue に対する詳しさを Local Expert をランキング形式で提示を行う。

本研究では、はまぞうに 2002 年 11 月 30 日から 2013 年 7 月 3 日の間に投稿された記事 5,059,894 件を用いる。また、F-Venue は、2013 年 10 月 7 日の時点までに登録された 10,918 件を用いる。

本実験において、Local Expert 抽出で扱う Venue は、Foursquare の Venue の categories が、“Food” または “Nightlife Spot” に分類されているもの (レストラン系) と、“Bed & Breakfast”, “Boarding House”, “Hostel”, “Hotel”, “Motel” または “Resort” となっているもの (ホ

テル系) を用いる。レストラン系の評価項目は、食べログ⁷、ロケタッチ グルメ⁸、ぐるなび⁹ に基づいて選出した 6 項目を適用し、また、ホテル系の評価項目は、楽天トラベル¹⁰、Yahoo トラベル¹¹、じゃらん¹² に基づいて選出した 6 項目を適用した。各評価項目、および各分類義語を表 2、表 3 に示す。類義語は、WordNet¹³ に基づいて、評価項目としてふさわしい単語を人手で選出した。表 2 と表 3 に示した評価項目と類義語は、3.2 節で述べた Local Expert 抽出に用いる。

表 4、表 5、表 6 に “オスピターレ” という Venue に対して、提案システムを用いて Local Expert 抽出を行い、ランキング形式で提示した結果を示す。表 4 は、2010 年 6 月から 2013 年 6 月の 3 年間 ($month = 36$)、表 5 は、2010 年 6 月から 2010 年 12 月の半年間 ($month = 6$)、表 6 は、2013 年 1 月から 2013 年 6 月の半年間 ($month = 6$) のブログに対して提案手法を適用した結果である。星マークが付いているのは、公式ブログとして検出したブログユーザである。

表 4 は、表 5 のランキング作成後に投稿された記事を考慮して新たにランキングを作成した結果である。表 4 において、時間の経過に伴い、増加した記事によってランキングに変動があることがわかる。例えば、表 5 には、表 4 のランキングに入っていなかった author “niimi” がランクインしていることがわかる。“niimi” の記事を見ると、2011 年 5 月 19 日から “オスピターレ” を会場として、ワインパーティーや、ワインセミナーを複数回開催し、そこで出てきたワインや、料理について言及した記事を投稿しているため、ランクインしたと考えられる。そのため、表 5 では、ランキングには出現せずに、表 4 では、2 位と評価されたと考えられる。

ここで、ランキングを作成するために用いるブログの期間による結果の変化について考察する。表 6 は、半年の期間の記事、表 4 は、3 年間の記事を利用している。

⁷<http://tabelog.com/>

⁸<http://gourmet.livedoor.com/>

⁹<http://www.gnavi.co.jp/>

¹⁰<http://travel.rakuten.co.jp/>

¹¹<http://travel.yahoo.co.jp/>

¹²<http://www.jalan.net/>

¹³<http://nlpwww.nict.go.jp/wn-ja/>

表 7 評価実験で使用した 5 つの Venue(レストラン系)

Venue	カテゴリ
オスピターレ	Italian Restaurant
魚魯魚魯	Japanese Restaurant
鳥造	Japanese Restaurant
弁いち	Japanese Restaurant
末広鮨	Sushi Restaurant

表 8 評価実験で使用した 5 つの Venue(ホテル系)

Venue	カテゴリ
はまきたプラザホテル	Hotel
ホテルウェルシーズン浜名湖	Hotel
ホテル九重	Hotel
館山寺サゴロイタルホテル	Hotel
浜名湖かんざんじ荘	Hotel

結果として、短期間である表 6 は、公式ブログがランクインせずに、author“officesola”がランクインしていることがわかる。“officesola”の記事を見ると、2013 年 5 月 15 日から“オスピターレ”を会場として、マネーセミナーを複数回開催し、そこで出てきた料理に対して記事を投稿しているため、ランクインしたと考えられる。これにより、長期間に渡る記事を利用してランキングを作成すると、公式ブログが上位にランクインしやすい傾向があり、短期間にすると、最近に書かれた記事が多いブログユーザのランクを高く評価する傾向があることがわかる。そのため、今後の課題として、ブログユーザが長期的な観点から Venue について知りたいのかなどを考慮してランキングの作成に用いる記事の期間を決定する必要があると考えられる。

5 評価実験

本研究では、提案手法について、2 つの評価を行う。1 つは、人手によるランキング結果と比較することで、提案手法によるランキング結果の順位の妥当性を評価する。もう 1 つは、そのランキング結果に含まれるブログユーザが、Local Expert かどうか人手による判断と比較することで、Local Expert として適切かを評価する。

5.1 実験条件

本実験では、以下の手順で選出した Venue の中から無作為に選出された 10 個の Venue を用いる。ここで、10 個の Venue の内訳は、5 個のレストラン系の Venue、5 個のホテル系の Venue である。はじめに、レストラン系とホテル系の Venue の中から、それぞれ総チェックイン回数(注目 Venue が含まれる記事の総数)が 50 回以上の Venue を選出した。これは、総チェックインが少な

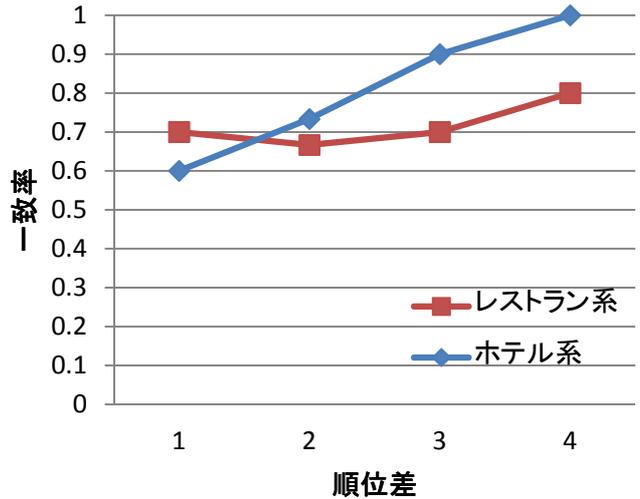


図 2 提案手法と人手によるランキング結果の比較

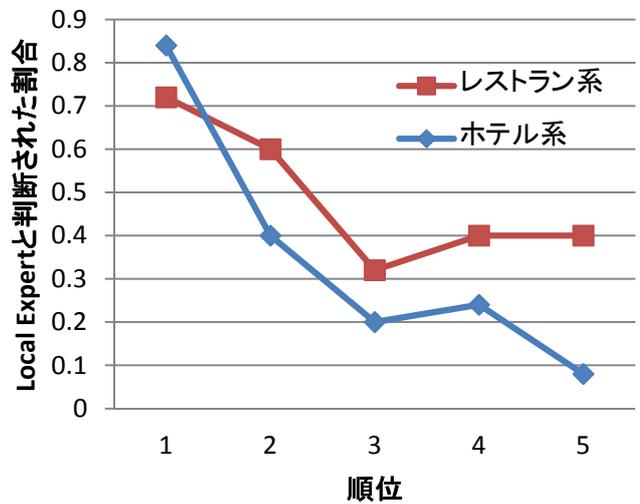


図 3 Local Expert 抽出の評価結果

い場合、ランキング結果に複数のブログユーザが入らない可能性があるためである。

5.2 ランキングの順位の妥当性の評価

本節では、提案手法によるランキングが Venue に詳しいブログユーザ順であるかを人手により評価する。評価方法は、提案手法により作成されたランキングの上位 5 人から選出した 2 人のブログユーザのどちらが Venue についてより詳しいかを人手により判断し、その結果が提案手法によるランキング結果の順位関係と一致しているかを評価した。

人手による評価は、実験協力者に上位 5 人のブログユーザの記述した Venue の単語を含むすべての記事を読んでもらい、2 人のブログユーザのどちらが詳しいかを実験協力者が判断した。結果として、5 人のブログユーザから 2 人のブログユーザのすべての組み合わせ(10 通り)について、どちらのブログユーザが Venue についてより詳しいかが得られた。最終的に、得られた 10 通りの 2 人のブログユーザの組み合わせにおいて、Venue に

対する詳しさの優劣と、提案手法のランキング結果が一致しているかを評価した。本実験に用いた Venue は、表 7、表 8 に示したものである。また、この評価の際は、 $month = 120$ とした。

レストラン系、ホテル系のカテゴリの順位差のそれぞれの平均値を図 2 に示す。図 2 において、順位差とは、提案手法によるランキング結果におけるブログユーザの順位差である。例えば、2 位のブログユーザと 5 位のブログユーザの順位差は 3 である。

図 2 において、どちらのカテゴリも順位差が大きくなれば正解率が上昇する傾向があることがわかる。そのため、提案手法は、順位差が大きいブログユーザについては正確にランク付けできていることがわかる。結果として、提案手法は、ブログユーザのランク付けを大きく間違えることは少ないと考えられる。

5.3 Local Expert 抽出の評価

本節では、提案手法によるランキング結果において、Local Expert であるブログユーザが上位に位置づけられているかを評価する。評価方法は、前節と同じ上位 5 人のブログユーザの中に Local Expert として相応しいブログユーザが何人含まれているかを評価する。実験協力者に上位 5 人のブログユーザの記述した Venue の単語を含むすべての記事を読んでもらい、そのブログユーザが Local Expert に相応しいかを判断してもらった。本実験に用いた Venue は、表 7、表 8 に示したものである。また、この評価の際は、 $month = 120$ とした。

レストラン系、ホテル系のカテゴリに対する結果を図 3 に示す。図 3 において、どちらのカテゴリも下位ほど Local Expert と判断された割合が減少していく。そのため、上位では、提案手法が有効に機能したため Local Expert が抽出できたと考えられる。一方、下位においては、Local Expert と判断された割合が減少しているのは、Venue によらずに上位 5 人を算出し、評価を行なったことが大きな原因の 1 つとして挙げられる。これは、1 つの Venue についての Local Expert の人数は、Venue によって大きく変わることが考えられるためである。例えば、人気のある店名の場合、その店名について言及するブログユーザの数は多いと考えられ、逆もまたしかりである。現段階の提案システムでは、どのような Venue に対しても 5 位までのランキング結果をユーザに提示している。そのため、Local Expert に相応しい人数によって、ユーザに提示する人数を調整する必要がある。これは、1 人のブログユーザの合計値が、全ブログユーザの合計値に対して、一定以上の割合を占めている場合に、Local Expert とすることで対応できると考えられる。

6 おわりに

本研究では、ブログサイトにおいてゲーミフィケーションをホットスタートさせるため、地域ブログサイト全量解析に対する Venue の詳しさに基づいたブログユーザのランキング手法を提案した。ランキングを作成する際に、情報の鮮度を考慮するため、投稿日時を利用して記事の重みを求めた。提案手法により、時間の経過に伴い、増加した記事によってランキングに変動があることが確認できた。また、ランキングを作成するために用いる記事の期間を変えることでも、ランキングに変動があることが確認できた。さらに、評価実験により、提案手法によるランキング結果は、ブログユーザのランク付けを大きく間違えることは少ないことや、ランキングの上位には Local Expert が検出できていることを示した。

今後の課題として、4 章で述べたように、式 (1) の $month$ の適切な値を調査することが挙げられる。また、公式ブログの検出に用いる閾値を調整することが挙げられる。さらに、レストラン系とホテル系以外のカテゴリの Venue に対する評価項目を決めることや、評価項目を自動で決める手法の検討が挙げられる。加えて、1 つの記事に複数の Venue が出現した場合に対処することが挙げられる。これは、北原らの研究 [9] を参考に、記事内で 1 つの Venue の影響が及ぶ範囲、および強さを測ることで対応できると考えられる。

参考文献

- [1] ジェイン・マクゴガニル：『幸せな未来は「ゲーム」が創る』、早川書房、2011。
- [2] 河野 愛樹、中村 健二、小柳 滋：マイクロブログから抽出した地物情報と投稿間隔を考慮した位置情報推定、情報処理学会第 73 回全国大会、2011。
- [3] 倉島 健、手塚 太郎、田中 克己：Blog からの街の話題抽出手法の提案、電子情報通信学会第 16 回データ工学ワークショップ、2005。
- [4] 中嶋 勇人、太田 学：旅行ブログ記事からの名所とその付随情報の抽出、第 5 回データ工学と情報マネジメントに関するフォーラム、2013。
- [5] 平山 拓央、湯本 高行、新居 学、高橋 豊：属性評価モデルに基づく商品評価の抽出と提示、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、2011。
- [6] 小西 卓哉、木村 文則、前田 亮：周辺文を考慮するトピックモデルを用いた評価側面の推定、第 4 回データ工学と情報マネジメントに関するフォーラム、2012。
- [7] Bao J, Zheng Y and Mokbel M. F.: Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data, ACM SIGSPATIAL GIS'12, 2012.
- [8] 稲垣 陽一、中島 伸介、張 建偉、中本 レン、桑原 雄：プロガ体験熟知度に基づくブログランキングシステムの開発および評価、情報処理学会論文誌：データベース、Vol.3, No.3, pp. 123-134, 2010。
- [9] 北原 沙緒理、田村 航弥、波多野 賢治：Web テキストにおける内容密度分布の抽出とその評価、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、2011。