3C1-1

# RDF データベース構築によるユーザの気分に応じた 観光スポット推薦システムの提案

Proposal of Sentiment-based Tourist Spot Recommendation System Using RDF Database

坂元 陽亮 \*1 高間 康史 \*1
Yosuke Sakamoto Yasufumi Takama

# \*1首都大学東京大学院システムデザイン研究科

Graduate School of System Design, Tokyo Metropolitan University

This paoer proposes a system that recommends tourlist spots around the current location relating with the sentiment selected by the user. This system collects the tweets and comments about tourist spots from Twitter and Yelp, and caluculates 10 kinds of emotional scores such as "joy" and "angry" for each spot. Those scores are stored in a database and used for spot recommendation. Effectiveness of the system is evaluated by applying it to tourist spot recommendation in Nakano (Tokyo, Japan).

# 1. はじめに

本稿では、ソーシャルメディアの分析結果に基づき、ユーザの気分に応じた観光スポットを推薦するシステムを提案する. 対象地域を中野区とし、Twitter と観光スポット情報サイト Yelp から観光スポットに関する感想を収集し、感情表現辞典 [1] の感情を参考に 10 種類の基本感情に関するスコアをスポットごとに算出する. また、RDF により観光スポットデータベースを構築し、ユーザが選択した気分にあった現在地近隣のスポットを推薦するシステムを提案する. 距離ベースのスポット推薦手法との比較を通じて、新規性の観点から評価、考察を行う.

# 2. 関連研究

# 2.1 テキストからの感情分析

テキストからの感情分析には、感情語辞書と対象となるテキストの単語あるいは、語句レベルでのマッチングによる手法が多く用いられる。感情の分類方法は、様々な方法が提案されている。山本らは Twitter 特有の感情の相関を考慮して、Plutchik の感情モデル [2] から次元を削減する方法を提案している [3].

# 2.2 観光スポット推薦

観光スポット推薦の手法は、ユーザの旅行履歴に基づき推薦するユーザベースの手法と、観光スポットの情報に基づき推薦するアイテムベースの手法の2つに大別される。本研究と同じアイテムベース手法として杉浦らは、旅行者にアンケートを行い、ユーザの観光スポットにおける感情を調査している。その結果に基づき観光スポットの推薦を行う手法を提案している[4].

# 3. 提案手法

#### 3.1 感情分析

本稿では、ユーザの気分に応じた推薦システムの作成のためにテキストから感情の抽出を行う. 感情表現辞典 [1] の 10 の基本感情に分類し、感想から各感情の値を計算する. 感情表現辞書に掲載されている単語や文節に加え、各単語の類義語を日本語 WordNet\*1 の類義語API を用いて収集し、合計 11,870 語の感情語辞書を構築した. 表 1 に各感情における単語の数と例を示す.

本稿では、中野区の観光スポットを対象に推薦システムを構築する。観光スポットは、中野区都市観光情報サイトの「まるっと中野」に掲載されている中野区認定観光資源  $^{*2}$  と、中野区逸品グランプリ  $2015^{*3}$  にエントリーをしている店舗の計 230 施設を使用する。TwitterAPI

連絡先: 高間康史, 首都大学東京大学院システムデザイン研究 科, 東京都日野市旭ヶ丘 6-6, Email: ytakama@tmu.ac.jp と観光情報サイト Yelp の YelpAPI\*4 を用いて感想を取得し、分析して観光スポット毎に感情スコアを計算する。 Twitter と Yelp の特性から情報の確実性を考え、Yelp のレビューが存在する観光スポットは Yelp のレビューから感情値を計算し、存在していないスポットは、Twitter の Tweet から感情値を計算する。感情値の計算には、データセットの各感想内にある感情語辞書内の単語の出現回数あるいは TF-IDF 法を用いる。

表 1: 各感情の感情語辞書への登録単語数

感情	登録単語数	単語例
哀	1,041	泣く、ジーンと、ポロポロ
安	497	寛ぐ,すっきり,やれやれ
厭	3,254	妬む,やきもち,忌み嫌う
喜	787	喜ばしい,歓喜 ,ニコニコ
驚	417	ぎょっとする、仰天、びっくり
好	1,396	惹かれる、陶酔、うっとり
恥	77	恥じる, 羞恥, 恥辱
怒	1,000	心外, 自棄, 憤る
怖	1,592	臆する、おどおど、びくびく
昂	1,809	感嘆,逆上せる,やけっぱち
計	11,870	

# 3.2 データベース構築

提案システムでは、データベースに RDF データベースを用いる.トリプルの主語には各観光スポットの情報ページの URL を使用し,目的語には各述語に対応した情報を URI またはテキスト形式で記述する.

SPARQL エンドポイントの構築にオープンソースソフトウェアの Fuseki\*5 を使用し、SPARQL クエリを使用することによって RDF データベースから必要なデータを JSON 形式で取得する。表 2 には、データベースのカテゴリごとのスポット数と、感情値が計算できたスポット数 (計算不可) を示す。

表 2: カテゴリ別の観光スポット数

カテゴリ	計算可	計算不可	計
グルメ	84	24	108
神社・仏閣	40	24	64
自然・公園	39	12	51
お祭り	4	3	7
計	169	61	230

# 3.3 推薦システム

推薦システムの構成を図1に示す.システムをAndroidのアプリとして実装し、スマートフォンを介してユーザへ情報を推薦する.ユーザにアプリを通じてその時の気分を入力してもらい、RDFデータベースにSPARQLクエリを送信し、入力された気分に合った観光スポットの情報を取得して推薦を行う.ユーザは、表3のような感情の組み合わせで表される、8種類の気分に関する選択肢から1つ以上選択する、また、利便性を考え、距離のオプションを設けた.図2に実際の

<sup>\*1</sup> http://nlpwww.nict.go.jp/wn-ja/

 $<sup>*2 \</sup>quad \text{http://www.visit.city-tokyo-nakano.jp/category/nintei}$ 

<sup>\*3</sup> http://www.nakano-no-ippin.com/

 $<sup>*4 \</sup>quad https://www.yelp.com/developers/documentation/v3$ 

<sup>\*5</sup> https://jena.apache.org/documentation

#### アプリ画面を示す.



図 1: 推薦システムの構成

表 3: 各気分を構成する感情

式 9. 日 X(力 と 情) X す る 心 旧				
気分	感情の組み合わせ			
心地よくなりたい	「喜」と「好」			
気分転換したい	「安」と「驚」			
写真を取りたい	「好」と「驚」			
静けさのある場所に行きたい	「怖」と「厭」			
心を落ち着けたい	「安」と「哀」			
驚く体験をしたい	「驚」と「昂」			
お腹が空いた	カテゴリがグルメのもののみ表示			
_ マイナースポットに行きたい	感情値の計算ができないマイナースポットのみ表示			

•	<b>♀</b> <sup>ug</sup> 2 9:02	₩ 0 17:
□ TFIDF重み □ カテゴリー重み	3km	江古田氷川神社
-つ以上気分を選択し、検索ボタンを押	してください	The state of the s
検索		NAME OF THE PARTY
心地よくなりたい	<u>~</u>	
気分転換したい		
お腹がすいた		
写真が撮りたい		カテゴリー
心を落ち着けたい		神社・仏閣 住所 東京都中野区江古田3-13-6
静けさのあるところに行きたい		京京都中野区江古田3-13-0 説明 江古田氷川神社は、寛正元年
驚く体験をしたい		(1460) 素戔嗚尊(すさのおのみこと)をここに祭ったことに始まると
マイナースポットに行きたい		伝えられています。 神楽殿は、弘化 4年(1847)に建立され、昭和7年
		(1932)まで本殿として使われていました。屋根は寄棟造りで、正面は
4 0		4 0 🗆

図 2: 実際のアプリ画面

## 4. 評価実験

# 4.1 実験内容

他の距離ベースのスポット検索サービス (GooglePlaceAPI\*6 と Yahoo ローカルサーチ API\*7) と、比較用に作成した距離ベースの 推薦システムにより生成される推薦リストと、提案システムによる推 薦リストを比較することで, 提案手法の有効性を新規性の観点から評 価する. 比較用に作成した距離ベースの推薦システムは、現在地とス ポットの座標から計算する直線距離の近い順に推薦を行うシステムで あり、推薦するスポットは提案システムと同じデータベースを利用す る. ユーザの気分を推薦に使用することによってどれだけ推薦の順序 が変化し、新規性の高い推薦が行えるか確かめるためにこの距離べて ス推薦システムを構築した. これらの距離ベースとの比較に新規性 (式 (1)[5]) を採用する. 推薦システムが推薦したアイテムリストのうち, プリミティブな推薦システムが推薦した推薦リスト (PM) と提案シス テムによる推薦リスト (RS) にどれだけ差があるかを示す評価指標で あり、値が大きいほど新規性が高いことを示す. 本稿ではプリミティ ブな推薦システムとして既存のスポット検索サービスや距離ベース推 薦システムを用いる.

新規性 = 
$$100 \cdot \frac{|PM \setminus RS|}{|RS|}$$
 (1)

#### 4.2 実験結果

表 4 は、提案システムを利用した時に推薦される推薦リストを全オプションの組み合わせについて作成し、それぞれで新規性を計算し、類似度と TF-IDF 重み、カテゴリ重みの設定ごとに平均値を求めた結果の一覧である. 結果を見ると、51%と高い値になっている. これよりユーザの気分を考慮することにより、新規性の高い推薦が行えていることがわかる. また、新規性の平均値は設定によらずほとんど同じであることがわかる. このことから、観光スポットの感情値自体がカテゴリの重みの割合に近いものが多く、かつ単語の出現回数は非常に少ないことが考えられる.

表 5 に、各気分における新規性を比較した結果を示す.この表は、気分 A の新規性から気分 B の新規性を引いた値を要素として持つ.推薦リストの計算の際にはオプションを指定せず、類似度は総和を用いている.「静けさのあるところに行きたい」の新規性が最も高いのは、これに対応する「怖」と「厭」の感情値がカテゴリ「神社・仏閣」と「自然・公園」で大きい傾向にあるため、「グルメ」のスポットが周辺に多い中野駅周辺で気分を考慮しない距離ベースで推薦を行った場合とは、推薦されるスポットの傾向が異なったことが原因といえる.また、「心地よくなりたい」と「写真が撮りたい」や、「心を落ち着けたい」と「気分転換をしたい」など、構成している感情が共通している程新規性の差が少ないことがわかる.このことから、感情の組み合わせを重複の無いように設定することで、気分ごとの新規性の差が大きくなり、推薦リストの多様性を高める効果が期待できる.

表 4: 中野駅における新規性の各平均値

2. 1 1 3 W ( - 60 ) 3 W ( ) W ( 2 ) U ( 1 ) U					
分類方法	オプション名または有無	平均値			
	コサイン類似度	51%			
類似度	ピアソン相関	52%			
	ユーグリッド距離	50%			
	総和	52%			
TFIDF 重み	有	51%			
	無	51%			
カテゴリ重み	有	51%			
	無	51%			
	平均	51%			

表 5: 気分ごとの新規性の比較

気分 A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) 心地よくなりたい	0						
(2) 気分転換したい	-4	0					
(3) 心を落ち着けたい	-5	-1	0				
<ul><li>(4) 写真が取りたい</li></ul>	-1	3	-1	0			
(5) お腹が空いた	-18	-14	-13	-17	0		
(6) 驚く体験をしたい	0	4	5	1	18	0	
(7) 静けさのある所に行きたい	1	5	6	2	19	1	0

#### 5. おわりに

本稿では、Twitter と Yelp を収集して感想の感情分析を行う手法の提案、観光スポットの RDF データベースの構築、ユーザの気分に応じた観光スポットの推薦システムの構築を行った。中野区の観光スポットを対象として既存の周辺スポット検索サービスや独自に実装した距離ベースの推薦システムとの比較実験を行った結果、提案システムにより新規性の高い推薦が可能であることを示した。

今後は、感情分析の精度向上や、ユーザ実験による推薦システムの評価、RDF データベースのオープンデータ化を行うことが考えられる.

## 参考文献

- [1] 中村明. 感情表現辞典. 東京堂 出版, 1993.
- [2] Plutchik, Robert. "1. A general psychoevolutionary theory of emotion," Theories of emotion, Academic Press, pp. 3-31, 1980.
- [3] 山本湧輝, 熊本忠彦, 灘本明代. ツイートの感情の関係に基づく Twitter 感情軸の決定, DEIM 2015, E5-2,2015.
- [4] 杉浦孔明. 観光スポット推薦アプリ「京のおすすめ」を用いた長期実証実験, 観光と情報, Vol. 10, No. 1, pp. 15-24, 2014.
- [5] Adamopoulos, Panagiotis, Tuzhilin, Alexander. "On unexpectedness in recommender systems: Or how to expect the unexpected," ACM Trans. Intelligent Systems and Technology, Vol. 5, No. 4, pp. 1-32, 2015.

<sup>\*6</sup> https://developers.google.com/places/

 $<sup>*7 \</sup>quad http://developer.yahoo.co.jp/webapi/map/openlocalplatform/v1/localsearch.html \\$