

# レビュー投稿者の不満を解消するホテル推薦システムの提案

林 裕次郎<sup>†</sup> 河合由起子<sup>†</sup> 中島 伸介<sup>†</sup>

<sup>†</sup> 京都産業大学 情報理工学部 〒 603-8555 京都府京都市北区上賀茂本山

E-mail: †{g2054065,kawai,nakajima}@cc.kyoto-su.ac.jp

**あらまし** ホテル検索システムにおけるレビュー文章は、ホテル検索を行う一般ユーザにとって有益である。しかし、レビュー投稿者に対するレビュー投稿のメリットは乏しく、レビュー投稿を行うユーザは多いとはいえない。そこで我々は、レビュー投稿者に対して有益な情報を提供可能なシステムを開発することで、レビュー投稿者を増加させることが出来るのではないかと考えた。本研究では、レビュー投稿者が投稿した不満レビューを分析し、この不満を解消するようなホテルをレビュー投稿者に推薦可能なシステムの提案を行う。

**キーワード** ホテル推薦, 情報推薦, レビュースコアリング

## 1. はじめに

近年、楽天トラベル [1] やじゃらん [2] など、宿泊施設に対する利用者のレビューや各評価項目に対するスコアが閲覧可能なホテル予約サイトが数多く運営されている。一般ユーザが、宿泊先のホテルを決定する際、他のユーザのレビューを参考にすることが多く、自身が宿泊したことがないホテルの情報を収集する手段として、このようなレビュー情報は非常に有益である。しかしながら、ホテルを利用したユーザのうち、レビュー投稿を行うユーザはごく一部に限られており、全てのユーザがレビュー投稿を行う理想的なケースに比べると、有効な情報が失われていると共に、限られたユーザの意見のみが参考にされるという意味でバイアスがかかる可能性が否定できない (図 1)。我々はこのような一部のユーザのみのレビューが共有される状況は好ましいことではないと考えている。

レビューを投稿しない理由としては、レビューを投稿しても他のユーザが参考になるだけであり、レビュー投稿者に対する大きなインセンティブは無いという理由が考えられる。レビュー投稿数増加を狙い、ポイント還元などを行っている場合もあるが、それはポイントという報酬を貰うためにレビュー投稿を行うこととなり、適切なレビューを投稿することよりも、質が悪くてもレビュー投稿数が増えれば良いという、悪い傾向の原因になりかねない。

そこで本研究では、ユーザにとってレビュー投稿自体にメリットがあれば、レビューを投稿するユーザが増加するのではないかと考えた。本稿では、ユーザが投稿した不満レビューに対して、これを解消するようなホテルをこのレビュー投稿者に推薦するシステムを提案する。

## 2. 関連研究

有名なホテル予約サイトとして楽天トラベルがある。楽天トラベルでは、ホテルや旅館など多くの宿泊施設がレビューと共に掲載されており、評価を比較しながらホテルの検討ができる。楽天トラベルのようなレビューを取り扱うサイトが広く普及したことで、レビューを分析する研究が多く行われている。

安倍らは、ホテルの評価項目を詳細化した評価表現辞書を作成し、自動スコアリングを行う方式を提案している [3]。

青柳らは、宿泊予約サイトのレビュー文からクレームを抽出し、内容解析を行い課題抽出する方法を提案している [4]。

林部らは、宿レビューから「肯定的事実」と「推薦対象」という言語知識を、宿のレビューから抽出する手法を提案している [5]。

田熊らは、ユーザが順位付けしたホテルの評価項目を元にホテルを推薦する手法を提案している。投稿されたレビューから、評価項目ごとの名詞の数をカウントし、数が多い順に評価項目の順位付けを行い、ユーザ類似度が高いレビューを推薦している [6]。

和田らは、宿泊予約サイトのレビューに記述されているギャップ文を抽出し、評価項目ごとに提示する手法を提案している [7]。

上述の通り、ユーザによって投稿されたレビューや特徴ベクトル、ランキングなどを分析し、情報推薦を行う研究が数多く行われている。これらはこれからホテルを予約しようとしている一般ユーザに対して有益な研究だが、レビュー投稿者にメリットを与えるような研究は行われていない。

## 3. レビュー投稿者の不満を解消するホテル推薦システム

### 3.1 提案システムの概要

提案システムの概要を図 2 に示す。ユーザがホテル宿泊後レビューを投稿する。投稿された不満レビューを評価表現辞書を用いてスコアリングしレビュースコアを作成する。同様の評価表現辞書を用いてホテルスコアも作成する。レビュースコアとホテルスコアから不満点を解消する不満解消スコアを作成する。その後不満解消スコアと一致度の高いホテルを検索しレビュー投稿者に推薦する。

### 3.2 ホテルデータベース

本節では、過去のレビューデータを保存する、ホテルレビューデータベースについて述べる。ホテルレビューデータベースには、過去にレビューを行ったユーザのデータが、図 3 のように格納されている。このデータベース内にあるデータを用いて、推



図1 一般的なレビュー投稿の流れ

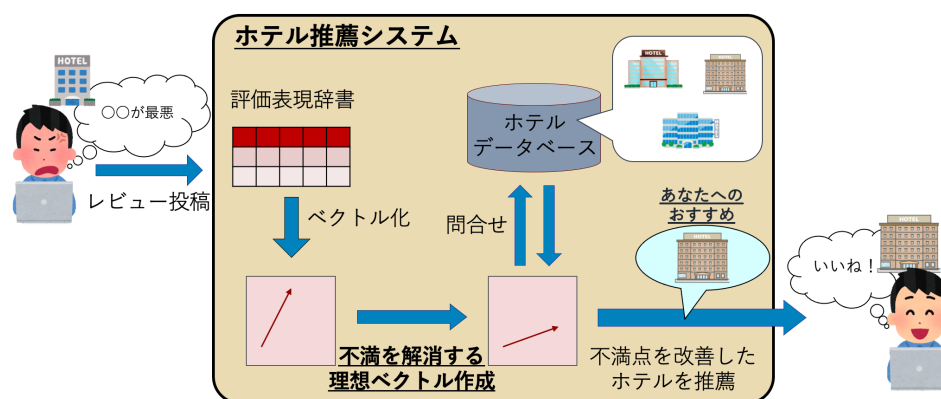


図2 レビュー投稿者の不満を解消するホテル推薦システムの概要図

薦を行う。このデータベースの管理にはデータベース管理システム MySQL を使用しており、各カラムの内容はユーザ判別を行うための "username"、ホテル名の "hotelname"、レビュー文が保存されている "review- text"、それ以外の "location"、"room"、"food"、"bath"、"service"、"facility"、"total" は各項目に対する評価という形になっている。このデータベースを作成するサンプルデータとして、楽天トラベルのレビューデータ約 250 万件を採用した。

### 3.3 評価表現辞書

本節では、先行研究で行われた、ホテル特化型評価表現辞書の作成方法について述べる。ユーザは様々な観点でホテルに対するレビューを投稿している。そこで、実際に使用されている表現を辞書に登録することで項目別評価表現辞書を構築した。辞書の構築に利用するホテルレビューデータはホテル予約サイトである楽天トラベルのレビューデータを採用した。具体的な辞書の構築は、過去に投稿されたレビュー文に含まれている評価表現をキーワード、特徴、程度に分割し評価項目ごとに分類して辞書に登録する。

表1に従来の評価項目と先行研究で提案した詳細化した評価項目の一覧を示す。これらの詳細化した評価項目に対する評価表現に対してスコアを設定し、評価表現辞書に登録する。評価表現に対してのスコア設定は、20代の男女5名にアンケートを行い0から5の範囲でスコアをつけてもらったものの平均と、

表1 従来の評価項目と提案する詳細化した評価項目

従来の評価項目	詳細化した評価項目
立地	コンビニ・スーパーへのアクセス 周辺の飲食店 観光スポットへのアクセス ホテルまでのアクセス
食事	豪華さ 品数 量 味
部屋	眺め 広さ 清潔さ 静かさ
風呂	温泉 眺め
サービス	スタッフ 送迎
設備・アメニティ	アメニティ ネット環境 機器

楽天トラベルのデータを元に設定している。

### 3.4 評価表現辞書に基づくスコアリング手法

本節では、先行研究で行われた、未知のレビューデータに対する自動スコアリングを行う手順を述べる。詳細化した評価項

username	hotelname	location	room	food	bath	service	facility	total	reviewtext
		4	3	0	3	3	3	4	とても清潔な部屋でした。周辺は箱崎～千代にかけ
		5	5	4	5	5	5	4	貸し切り風呂がとて良かったです！色んな風呂が
		5	5	0	3	5	5	5	ダブルでの利用は初めてでした。相変わらず綺麗なリ
		4	4	0	3	4	4	4	フロントの方の対応も感じよく、部屋も清潔でした。
		2	5	4	4	3	4	4	部屋は広く、ベランダがあったのはとてもよかった。
		5	5	2	5	5	3	5	ホテルの外観も内装もとても素敵でした。部屋は広
		3	4	2	4	3	3	4	とにかく安いプランがあって部屋は広くて安くて良
		5	3	3	4	3	2	3	ビジネスホテルのユニットバスではなくつるがないの
		5	5	5	4	5	4	4	この度は両親子供と4人でお世話になりました。部
		5	5	4	3	4	4	5	駅から近くとても便利で良かったです。部屋もきれい
		4	4	3	3	4	4	4	いつも利用させて頂いています。今回はいつもと違い
		5	5	5	4	4	5	5	深夜に予約・チェックインしましたが、対応が素早
		4	4	4	3	5	2	4	夜遅くのチェックインだったのにも関わらず、嫌な
		4	4	5	4	4	4	4	久しぶりの利用でした。朝食が以前と比べて、凄
		5	4	0	3	4	3	4	二度目の宿泊です。施設は隅々まで綺麗にされてお
		4	4	5	4	5	4	5	多少、古さを感じましたが、清潔感もあり、芦ノ湖
		5	4	4	5	4	3	5	熊本出張でいつも「熊本」を利用しているので、
		1	1	0	1	1	1	1	二度と泊りません。【ご利用の宿泊プラン】【室数
		5	5	5	5	5	5	5	社員が宿泊させていただきましたが、感想として、
		2	3	0	2	3	3	3	値段相応の施設でした。単に寝て泊まる方には申し
		4	4	0	3	5	4	5	サービス人員真の非常親切、交流讓人愉快。旅館就在岡
		3	3	0	4	4	4	4	普通の都市圏でこの値段は十分過ぎるかと思います。
		4	4	3	3	5	5	4	同僚と泊まりましたが、満足してました。カウンタ
		4	5	5	4	4	4	5	今回初めて利用させて頂きました。他宿泊施設の一

図 3 データベース内データの一例

## 1. レビューに含まれている評価表現を辞書から探索

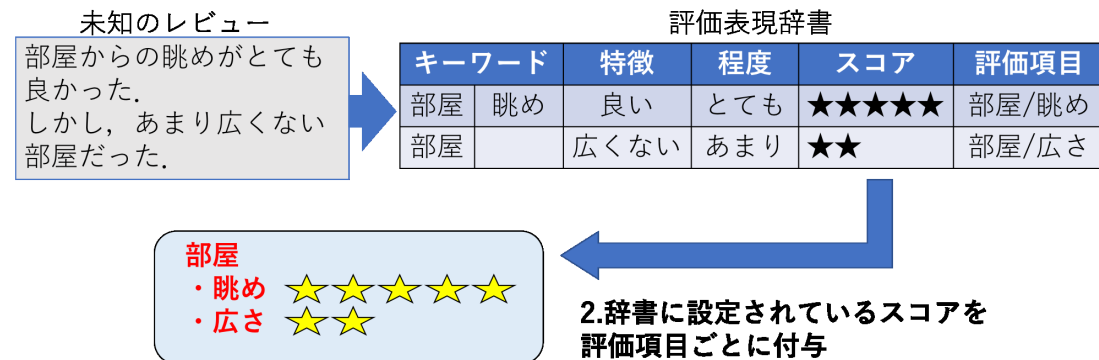


図 4 評価表現辞書を用いた項目別スコアリングの例

目ごとにスコアを付与するために、レビュー文に対して形態素解析エンジン mecab を用いて形態素解析を行う。分割されたキーワード、特徴、程度の組み合わせを評価表現辞書から探索を行い、辞書に当てはまる評価表現があれば、設定されているスコアをその評価項目に付与する。実際にスコアリングを行う過程を図 4 に示す。このレビューには、部屋の眺めに対する評価と部屋の広さに対する評価が書かれている。レビューを解析し、評価表現辞書に当てはまる表現があるので、設定されているスコアを部屋の「眺め」と部屋の「広さ」に付与している。

### 3.5 各ユーザの不満を解消するホテルの推薦手法

3.4 節にて説明した自動スコアリング手法を用いたレビュー投稿者の不満を解消するホテルの推薦手法について説明する。3.4 節の自動スコアリング手法を用いてレビュースコアとホテルスコアを算出する。レビュースコアとはユーザが宿泊したホテルに対してつけたレビューをスコア化したものである。ホテルスコアとはホテルデータベース内のホテルに付与されている全レビューをスコア化し、平均値をとったものである。また、本研究において、レビュー投稿者の不満とはレビュースコアにおいて、スコアが 2.0 以下の評価項目とする。スコアが 2.0 以下の評価項目が複数あった場合は全て不満点として処理する。レ

ビュースコアとホテルスコアから不満解消スコアを算出し、算出された範囲に当てはまるホテルをユーザに対して推薦する。

#### 3.5.1 不満解消スコアの算出方法

本節では、不満解消スコアの算出方法について述べる。不満解消スコアは、レビュースコアの不満点、レビューで言及のなかった評価項目に分類する。レビュースコアの不満点はホテルスコアと最大値との中間点以上を許容範囲とし、レビューで言及のなかった評価項目については宿泊したホテルのホテルスコア以上とする。これらの具体的な値は今後評価実験を用いて最適化していく。

#### 3.5.2 推薦順位の決定

本節では、3.5.1 節で決定された範囲内のホテルの推薦順位について述べる。推薦順位を決定する手法は現在 2 つ考えており、一つ目は「範囲内のホテルの内、総合評価の高い順に推薦する」、二つ目は「範囲内のホテルの内、不満点の評価項目のスコアが高い順に推薦する」である。これは今後の評価実験を通して決定する。

## 4. ま と め

本稿では、レビュー投稿者の不満を解消するホテル推薦システムについて提案した。また関連研究を調べ、本研究の新規性

を確認した。今後はこれらのシステムの開発に取り組み、評価実験を行う予定である。

## 謝 辞

本研究の一部は、科研費(課題番号:19K12243, 20H04293, 22H03700, 22K12281) および京都産業大学先端科学技術研究所(ヒューマン・マシン・データ共生科学研究センター) 共同研究プロジェクト(M2001)の助成を受けたものである。ここに記して謝意を表す。

## 文 献

- [1] 楽天トラベル <https://travel.rakuten.co.jp>
- [2] じゃらん net, <https://www.jalan.net>
- [3] 安部克, 中島伸介. レビュー自動スコアリング方式に基づくホテル推薦システム, DEIM Forum 2020, P1-28, 2020.
- [4] 青柳直人, 蓮池隆. 宿泊予約サイトのユーザーレビューにおけるクレーム抽出と内容解析による課題抽出, FIT2021, CE-002.
- [5] 林部祐太. 宿レビューからの肯定的事実と推薦の対象の抽出, 言語処理学会 第 25 回年次大会, 2019.
- [6] 田熊浩二, 福原楓, 亀井清華, 藤田聡. 口コミを用いた嗜好抽出に基づくホテル推薦手法, 情報処理学会第 78 回全国大会, 2016.
- [7] 和田知亜紀, 豊田哲也, 大原剛三. 期待と現実のギャップに着目した宿泊施設レビューの解析とその検索への応用, 情報処理学会, 2018.