

## ユーザーレビューを用いた宿泊施設の情報補完手法 A Method of supplementation with information on accommodations by user review.

辻井 康一  
Koichi Tsujii

津田 和彦†  
Kazuhiko Tsuda

### 1. はじめに

インターネット社会の進化により、宿泊予約はネット予約が主流となる時代になっている。その際、宿泊施設(以下、施設という)の予約判断材料の中で、利用者から投稿されるユーザーレビューが重要視されている[1]。

宿泊日	年齢 性別	同伴者	総合評価	料金	設備	サービス	立地	清潔度	お食事	お風呂
2011/02/26	20代 女性	友達	5	5	5	4	5	5	-	5

◎お客様の感想  
初めて利用させていただきました。お部屋をショコラルームにランクアップしてもらい景色もよく、とても感動しました。お風呂が独立で広く、洗面所、トイレも使いやすくてよかったです。最初に案内してくれた従業員の方が不慣れと言いかげ把握してないのか不安を覚えました。その他十分満足でした。また利用したいです。

図1 レビュー情報

一般的にユーザーレビューは、宿泊利用者が主要項目の数値評価と感想コメントをそれぞれ書き込み、投稿順に並べた固定的な表示形式である。ビジネス、観光と旅行目的が異なる場合、宿泊検討者が知りたい情報、宿泊利用者が何に満足を感じたかの内容は異なる[2][3]。しかしながら、その違いをわかりやすく表現できているとは言い難い。そこで本研究では、レビューを分析し、旅行目的別の評価項目の具体的な特徴の発見し、施設予約時に宿泊検討者が知りたい情報を提供するシステムの構築を目指す。

### 2. 投稿の概要

部屋タイプ別のレビュー投稿件数と、その基礎的属性を表1に示す。期間は2009年1年間に、施設の代表的な客室タイプであるシングルルーム、ツインルーム、和室を利用した宿泊利用者から投稿されたデータを基とした。性別は宿泊代表者の性別である。

表1 部屋タイプ別の評価平均と投稿件数

	男性	女性	件数計	比率	平均文字数
データ数	6,126	2,055	8,181		66
シングル	4,028	946	4,974	61%	58
ツイン	905	614	1,519	19%	72
和室タイプ	1,193	495	1,688	21%	83

なお、シングルルームの宿泊利用者は主にビジネス利用であり、またツイン、和室の利用者は観光利用がほとんどであるため、明確な特徴が出やすいと想定される。

### 3. 宿泊施設の特徴情報抽出

#### 3.1 基礎データ数と数値評価による傾向

大半の予約サイトにおける数値評価は、表2に示す7種が規定されており、宿泊利用者はそれぞれ評点を付けて投稿する。某サイトの投稿者の評点の平均を表2に示す。シングル、ツイン利用者では「設備」「清潔度」の評価が高く、

「サービス」「お風呂」の評価が低い。和室利用者では「設備」「お風呂」の評価が高く、「サービス」「お食事」の評価が低いことが分かる。これらは、旅行目的による注目項目の違いを表わしていると考えられる。これら数値評価の高低の理由が感想コメントに書かれている可能性が高い。そのため、感想コメントより注目される評価項目のキーワードを抽出し、それがどのような特徴と関係するのかを分析する。

表2 部屋タイプ別の数値評価の平均値

	シングル	ツイン	和室	全体
料金	3.80	3.97	3.80	3.83
設備	4.02	4.15	3.88	4.01
サービス	3.61	3.81	3.41	3.60
立地	3.72	3.92	3.80	3.77
清潔度	4.03	4.14	3.86	4.01
お食事	3.81	3.97	3.59	3.79
お風呂	3.32	3.73	3.88	3.56
総合評価	3.36	3.51	3.77	3.47
投稿件数	4,974	1,519	1,688	8,181

### 3.2 テキストマイニング分析

数値評価によって導き出された注目項目に着目し、感想コメントを分析する。テキストマイニングツールは、数理システムのText Mining Studioを利用した。

感想コメントのなかにも出現する特徴表現を分析する。まず特徴的な単語を抽出し、投稿者が注目する内容の傾向を把握し、特徴語として分類していく。次にその特徴語が、どのような評価を受けているのかを検討するため、分析ツールの特徴表現抽出機能を用い、利用者が注目する特徴語がどのような表現で記載されているかを係り受け解析により特徴表現を分析する。

## 4. ユーザーレビューの分析

### 4.1 特徴語の分類

類義語等の辞書登録の後、マイニングを実施する。その際、全体における単語頻度の高い特徴語を表3に示す。客室の評価である「部屋」、「風呂」、「宿泊施設」、「従業員」の順に出現頻度が高く、前述のユーザーレビューの数値評価項目に関係が深い単語の出現が多くみられる。

表3 特徴語の分類と頻度、割合

	品詞	頻度	割合(%)
部屋	名詞	2,030	2.5
風呂	名詞	1,722	2.2
宿泊施設	名詞	1,523	1.9
従業員	名詞	1,360	1.7
利用	名詞	1,146	1.4
満足	名詞	997	1.2
朝食	名詞	962	1.2
清潔	名詞	956	1.2
立地	名詞	946	1.2
利用+したい	名詞	904	1.1

## 4.2 特徴表現分析

特徴表現の評価は、感想テキスト全体での係り受け特徴表現の頻度である全体頻度と、部屋タイプを属性とし、その属性内での特徴表現の出現頻度である属性頻度を用いて比較する。また指標値とは、その属性の中で特徴的である係り受け表現の頻度を補完類似度により計算したものである。なお、補完類似度の計算式は以下のとおりである。

本稿では誌面の関係上、ビジネス目的でも観光目的でも注目される風呂、食事、立地に関する結果を記載する。

補完類似度の定義：
$$\frac{ad-bc}{\sqrt{(a+c)(b+d)}}$$

	属性j	属性j以外
特徴表現i	a	b
特徴表現i以外	c	d

### 4.2.1 特徴表現分析 =風呂=

特徴語として頻度の高い単語の一つである、風呂に対象を絞り込んだ係り受けによる特徴表現を、部屋タイプ属性別に分類し、指標値の順に並べた結果を表4に示す。

表4 風呂の特徴表現と頻度と指標値

	係り受け	属性頻度	全体頻度	指標値
シングル	風呂-狭い	54	91	38.75
	風呂-広い	30	64	10.87
	風呂-小さい	10	17	7.06
	広い-風呂	8	12	6.92
	風呂-有り難い	6	8	5.98
ツイン	大きい-風呂	3	8	6.68
	風呂-広い	9	64	5.85
	風呂設備-良い	2	5	4.57
	風呂-嬉しい	3	16	3.84
	良い-風呂	2	9	3.15
和室	温泉-良い	20	35	17.40
	風呂-残念	13	18	14.69
	貸切風呂-良い	8	8	11.23
	風呂-ぬるい	9	11	11.21
	温泉-気持ち良い	5	8	4.88

シングルルームでは、風呂の広さに関する表現が多い。「風呂-狭い、小さい」は一般的なビジネスホテルのユニットバスの窮屈さに対する指摘である。また大浴場が併設されるビジネスホテルなどに関しては「風呂-広い、有り難い」と表現されている。ツインにおいても広さが注目されるが、風呂自体の大きさに満足している。和室の場合は温泉、大浴場に関する感想となっている。温泉に対する感想、風呂に対する心配り、工夫の有無が注目されている。

### 4.2.2 特徴表現分析 =食事=

次に、食事を対象に分析した結果を表5に示す。

シングル・ツインの場合は朝食に関する記載で占められており、特に朝食バイキングの評価が多い。一方、和室においては、夕食に関する内容が記載されていることが特徴的である。食事について、夕食が宿泊サービスと一体化している和室と、食事は付加サービスであるホテルとは表現の違いが明らかである。

### 4.2.3 特徴表現分析 =立地=

次に、立地を対象に分析した結果を表6に示す。

シングルの場合、交通の利便性の高さ、静かさに加え、周辺の居酒屋、コンビニといった施設周辺店舗の便利さについて言及されている。ツイン、和室など観光利用の施設の場合は、交通機関との利便性、観光地へのアクセスの観点から評価されている。また施設公式サイトに掲載された最寄駅からの所要時間が実時間より短く書かれていた場合、その点の不満を指摘する記述がみられた。

表5 食事の特徴表現と指標値

	係り受け	属性頻度	全体頻度	指標値
シングル	朝食-良い	44	59	52.79
	朝食-美味しい	50	92	42.62
	朝食バイキング-美味し	7	9	8.67
	朝食-欲しい	6	10	5.84
	バイキング形式-良い	3	3	4.31
ツイン	朝食-美味しい	17	92	12.01
	バイキング-美味しい	5	17	7.58
	朝食-安い	3	6	6.24
	朝食-高い	3	6	6.24
	バイキング-良い	4	15	5.50
和室	食事-美味しい	50	84	38.89
	料理-美味しい	30	40	31.33
	料理-良い	26	34	27.67
	食事-良い	29	55	17.73
	お料理-美味しい	10	12	11.47

表6 立地の特徴表現と指標値

	係り受け	属性頻度	全体頻度	指標値
シングル	立地-便利	33	43	15.35
	立地-良い	173	282	12.33
	静か-立地	7	8	4.61
	立地-少い	4	4	3.32
	立地-問題+ない	4	5	2.12
ツイン	立地-悪い	3	6	6.85
	立地-良い	37	282	6.29
	大変便利-立地	2	4	4.57
	近い-立地	3	13	4.22
	立地-大変満足	2	6	3.82
和室	良い-立地	5	19	11.14
	近い-立地	3	13	6.17
	立地的-良い	2	4	5.61
	楽しい-立地	1	1	3.12
	弱い-立地	1	1	3.12

## 4.3 ディスカッション

利用者の数値評価に高低のある項目において、感想コメントのテキスト分析を行い、部屋タイプ別に評価項目の特徴表現に違いがあることが明らかとなった。旅行目的の違いと感想コメント内容は密接につながっており、数値評価で規定されている7種の評価指標では、宿泊検討者が期待する内容が十分に表されていないと言える。

また、投稿される感想コメントは宿泊利用者が実際に経験しないとわからない事柄であり、また宿泊施設が公式に提供する情報には書かれていない有益な情報であるといえる。この特徴を生かし、旅行目的に応じた付加情報を動的に表示させることで、宿泊検討者の施設判断により役立つことができると考える。

## 5. おわりに

今後は、より具体的に正確な旅行目的別の傾向の把握と特徴情報の掲載方法を検討し、宿泊検討者に対しリアルタイムな情報掲載手法の確立を目指す。例えば、規定されている7種の数値評価指標を、ユーザの旅行目的に沿って動的に変更するなどを検討する。

また、これらは利用者の生の意見であり、これらの内容を旅行目的にあわせて特徴表現を好評、不評に分けて精査し、宿泊施設情報に反映させることで、宿泊施設への改善や公式サイト情報の修正を促す材料として活用することも検討する。

## 参考文献

- [1] 長島 広太, “サービス商品の選択・評価におけるクチコミの機能”, 経営論集, 74号 (2009).
- [2] 森川 毅, “ホテル・旅館の評価: 心理学的側面からの分析”, 長崎国際大学論叢(2001)
- [3] 松尾 睦, 奥瀬 喜之, プラート・カロラス, “サービス・クオリティ次元に関する実証研究-SERVQUALの再検討-“, 流通研究 (2001)