

# 旅行のクチコミサイトにおける 利用者の必要性に応じた情報表示

石川 瑛子<sup>†</sup>      中山 伸一<sup>‡</sup>      真栄城 哲也<sup>‡</sup>

<sup>†</sup> 筑波大学 図書館情報専門学群

<sup>‡</sup> 筑波大学 図書館情報メディア研究科

## 1 はじめに

インターネット上のクチコミサイトには、投稿者が付ける評価点に応じて、扱っている商品をランキング表示する機能がある。本研究は、利用者が必要とする情報の重要度に基づいてランキング表示する手法を提案する。ここではホテル等の宿泊施設についてのクチコミサイトを対象とする。

本研究の目的は、インターネット上のクチコミサイトに寄せられる様々な意見を利用者の視点から分類・表示した、既存のクチコミサイトよりも利用者が必要とする情報の重視度を考慮した表示方法の提案である。クチコミサイトには家電製品、化粧品や食品など様々なものがあるが、本研究では宿のクチコミサイトを対象とする。

既存の宿のクチコミサイトでは、クチコミが投稿日の新しいものから順に表示されており、利用者は自分の知りたい項目についてのクチコミを1つずつ読んで探していかなければならないため、必要としている情報を得るのに時間がかかってしまう。一方、クチコミの投稿者は、文章以外に宿を点数評価することも可能である。その際、総合点数だけでなく、クチコミサイト毎に差異はあるが、「食事」、「部屋」、「接客」、「清潔感」、「風呂」等の項目について点数評価できる。このため、総合点数の高い順や、特定の1項目に限ったランキング表示も可能である。しかし、既存のクチコミサイトには、それらの項目を同じように重視したものか、1つの項目についてのランキング表示しか存在しない。

このような表示方法では、多様化しているユーザーの好みに対応できず、さらに宿に対して重視する項目も人によって異なる。このことから、個人の重視度に合ったランキングを出す必要がある。本研究で

は、これらの問題を解決するために、クチコミをその種類毎に分類して表示し、利用者が必要とする情報を見付けやすくする方法について検討した。また、利用者に評価点の項目をどの程度重視するのかを設定してもらい、それに基づいてランキングを算出する方法についても検証した。

## 2 方法

まず、温泉宿を決める際に考慮する重視度を項目毎に0から100の値で記入してもらい、設定した数値に理由がある場合には記入もらう。数値は高い程、重視することを意味する。調査項目は、サービス、立地、部屋、設備、風呂、食事の6項目である。調査会社による宿を予約する際に重視する項目の調査結果も存在するが、平均的なデータではなく、個人の重視度についてのデータにより価値があるため、このような調査を行った。

旅行の予約サイト楽天トラベル<sup>1</sup>に掲載されている温泉宿を5軒選択し、それぞれの宿について最近の100件のクチコミに記載されている評価点数およびクチコミの文章から、調査に用いた6項目についての評価点を算出した。評価点の偏りを避けるためクチコミの投稿数が100以上の温泉宿を選択した。文章の内容からの抽出は、文章中に6項目のどれかについて記述されていれば、肯定的か否定的な内容かで1点または3点を加算する。どちらの意見も記載されているか、平均的だとする意見であれば、2点とする。なお、文章中に記載が無い項目については、クチコミの投稿者の意見が不明であるため、点数を付与しない。このようにして、対象とする宿について、クチコミから投稿者が付与した評価点と、入力した文章それぞれから2通りの平均点を、設定した6項目それぞれに算出できる。また、全項目の平均点が総評価点となる。

<sup>1</sup><http://travel.rakuten.co.jp/>

<sup>1</sup>Information Presentation for Travel Reservation Web Sites according to the User Needs

<sup>2</sup>Eiko Ishikawa, University of Tsukuba

<sup>2</sup>Shin-ichi Nakayama, University of Tsukuba

<sup>2</sup>Tetsuya Maeshiro, University of Tsukuba

このデータを用いて、個人毎に設定した 6 項目の重視度に基づき温泉宿の順位付け (ランキング) を行う。

評価項目毎に以下の式で評価点を計算する [1].

$$E_i = \log_{10}(N) \times w_i \times H_i \times \frac{H_i}{L_i} \quad (1)$$

なお、 $i$  は評価項目、 $E_i$  は評価項目  $i$  の評価点、 $N$  はその宿についてのクチコミの全投稿数、 $w_i$  は評価項目  $i$  についてのユーザの重視度、 $H_i$  は評価項目  $i$  について肯定的または高評価なクチコミの全投稿数  $N$  に対する割合、そして  $L_i$  は評価項目  $i$  について否定的または低評価なクチコミの全投稿数  $N$  に対する割合である。

総評価点は、

$$E = \sum_i E_i \quad (2)$$

で与えられる。

式 1 は化粧品のカチコミサイトを対象にランキングを計算するために検証し、その有効性を確認している [1].

このように、5 つの温泉宿について、サービス、立地、部屋、設備、風呂、食事の 6 項目の評価点を計算し、さらに平均点を求める。このように算出した点数に基づいて順番付けした温泉宿のリストと、元々の表示順序に並べたリストを被験者に提示し、どちらのリストが良いと思うかと、温泉宿の表示順序が自分の希望に沿っているかを、温泉宿の紹介資料等を見て判断してもらう。

### 3 結果および考察

大学生 10 人に、温泉宿を決める際に考慮する重視度を項目毎に 0 から 100 の値で記入してもらった (表 1)。このように、重視する項目は個人差が大きい。

表 1: 項目毎の重視度、被験者 2 人の例

被験者	サービス	立地	部屋	設備	風呂	食事
A	70	60	50	30	85	60
B	80	80	80	60	0	100

本研究では、宿の評価点を、クチコミ投稿者が入力する評価点と、記入する文章から抽出した評価点の 2 種類の数値から算出している。解析の結果、これら 2 つの数値は相関が低く、一致しないことが明らかになった。このことは、利用者が投稿する際に、文章で

記述する感想と、数値で付与する点数は必ずしも一致せず、異なる判断基準または思考過程によって利用者が行動していると推測できる。他の宿に対して同様の分析を行ったが、類似の結果が得られた。その場合、相関係数は  $-0.06$  と無相関であった。そのため、評価点は投稿文の内容から算出した値を用いた。

項目毎の評価点は式 1 を用いて計算できる。この値から、5 軒の温泉宿の表示順序が決定される。表 2 から、ランキングの順序は変化することが明らかである。

表 2: 項目の重視度を考慮した文中からの評価点によるランキングの変化。サイト掲載はクチコミサイトに表示されている評価点に基づく順位。被験者 A と B はそれぞれの期待値の重みを考慮した順位

サイト掲載	被験者 A	被験者 B
1	1	2
2	5	5
3	3	3
4	4	4
5	2	1

投稿者が入力する評価点は、入力可能な数値の範囲が 1~5 のように比較的広い場合でも、異なる宿の間で点数の差が出にくい問題がある。また、大数の法則のような統計的な力学が働くことも差が減少する要因の 1 つだと考えられる。このような値の小さい差はランキング表示の際に順番が隣接する宿や商品の表示が有効に活用できないことを意味する。さらに、投稿者が入力する数値に基づいてランキング表示がされているが、より正確なランキング表示には、クチコミ内容から評価点を算出する必要がある。文章からキーワードを抽出する方法は情報検索で活発に研究されているが、より精度の高い手法が必要である。本研究のように、利用者の好みや重視度を考慮することで、より利用価値の高いランキング表示が可能となる。また、本手法は、温泉宿だけでなく、他の種類の宿泊施設や商品についても同様に使える。

### 参考文献

- [1] 色摩, 中山, 真栄城, 「化粧品のクチコミサイトにおける期待を考慮したランキング表示」, 2009 年電子情報通信学会, 2009.