

楽天トラベルの口コミ情報の定量化と ホテル推薦のパーソナライゼーションの試み

福原 楓 亀井 清華 藤田 聡
(広島大学 大学院工学研究科 情報工学専攻)

概要 ホテル比較サイトにおけるホテルのレーティングを各利用者の重視項目によってパーソナライズ(個人化)する手法を開発した。重視項目は利用者が投稿した口コミ情報をもとに定量化される。この手法によって、自分と似た嗜好をもつ利用者が高く評価したホテルがより適切に推薦されるようになる。

◎ 背景と研究目的

楽天トラベルのようなホテル比較サイトでは、アンケート結果(☆の数)の単純平均がホテルのレーティングを決めている。しかし、たとえば評価者がアメニティと部屋の広さのどちらを重視したのかなどによって☆の意味合いは変わるはずである。

本研究の目的は、アンケートによるホテルのレーティングを利用者の嗜好にあわせて**パーソナライズ**するための新しい手法を開発することである。

◎ ホテルスコア

仮説1：利用者にとってこだわりのある項目の評価結果は、そうでない項目の評価結果よりも信憑性が高い。
仮説2：こだわりのある項目は、その利用者が投稿した口コミ情報のなかで頻繁に言及されている。



初めて利用しました。
快適なホテルでした。
スタッフの方の対応が良かったですが、
立地が不便でした。

項目名	評価
サービス	☆☆☆☆
立地	☆
部屋	☆☆☆
設備・アメニティ	☆☆
風呂	☆☆☆☆☆
食事	☆☆☆
総合評価	☆☆☆

キーワードの出現を項目毎にカウントし、ランキング

キーワード数は631(楽天の口コミから手作業で抽出)

[1, 2, 4, 6, 5, 3] … サービスを最重視。立地は二番目。
※上位2項目が同率のとき[1,2, …]と[2,1, …]の二つを生成

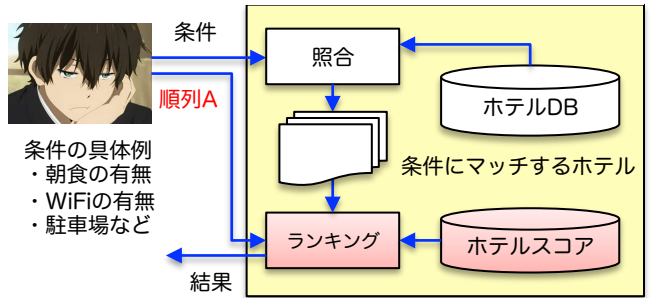
- ・ 検索者は全項目に順序付けをする(順列A)
- ・ 順列Aと一致する順列をもつ投稿者の評価を、それ以外の投稿者の評価の50倍重視
- ・ 順列A中のj番目に重視する項目を(7-j)倍する



→ 検索者の嗜好と投稿者とのマッチングを反映したホテルスコア

氷菓 © 米澤穂信, 京都アニメーション

◎ 提案するホテル推薦システム



◎ 評価

楽天トラベルにある広島市内のホテル・旅館72件と、約7万件の口コミ情報から、ホテルスコアを計算。提案方式と、提案手法からホテルスコアを除いた手法(楽天方式)、提案手法から条件照合を除いた手法(嗜好のみ)の三方式を比較。

三つの方式で上位3位に選ばれた合計(高々)9件のホテルを提示し、利用したいホテル上位3件を選択してもらおう(15名の被験者にそれぞれ5回ずつ試行してもらおう…合計75)。

選択	提案方式		
	1位	2位	3位
1位	41	13	11
2位	13	22	7
3位	8	16	22

選択	楽天方式		
	1位	2位	3位
1位	37	8	8
2位	14	19	9
3位	11	14	8

提案方式は被験者の好みにあったホテルを推薦できている。楽天方式と比べて、被験者が1位としたホテルが1位にランク付けされた件数は10%、3位以内にランク付けされた件数は20%それぞれ多い。

ランキングのマッチングを次のように数値化する：被験者が選んだ1位に5点、2位に4点、3位に3点、各方式が選んだ1位に3点、2位に2点、3位に1点をそれぞれ与え、あるホテルが両方にランクインしているとき、**点数の積をマッチングのポイントに加算する**(被験者が選んだ2位と方式が選んだ2位が同じとき、加算されるポイントは $4 \times 2 = 8$ となる)

提案方式	楽天方式	嗜好のみ
1420	1259	815

提案方式のマッチングのポイントは楽天方式よりも10%多い

◎ 今後の展開

検索者と投稿者の興味の一致の度合いを考慮することでホテルスコアの精度を向上させる(現在は、順列の一致のみで判断)。また自治体と協力するなどしてβ版を多くの人に使ってもらい、従来システムに対する優位性を検証する。