

アイテムのジオタギングに基づく 観光スポットベースクロスドメイン推薦システム

北村 統太[†] 奥 健太[†]

[†] 龍谷大学理工学部情報メディア学科 〒520-2194 滋賀県大津市瀬田大江町横谷1番5

E-mail: [†]t130464@mail.ryukoku.ac.jp, ^{††}okukenta@rins.ryukoku.ac.jp

あらまし クロスドメイン推薦システムは映画や音楽、本などの異なったドメインに跨って、ユーザにアイテムを推薦するシステムである。異なったドメインにおいて、共通アイテムや共通ユーザを手掛かりに推薦アイテムを決定する。しかし、共通アイテムや共通ユーザが存在しない場合には、推薦アイテムの決定は困難である。本研究では、共通知識として位置情報に着目する。例えば、ある観光地が舞台になった小説や映画があった場合、その観光地を共通知識としてクロスドメイン推薦に活用できると考える。本研究では、クラウドソーシングにより映画、音楽、本の内容に関連する位置情報を収集した。収集したデータは各ドメイン 10,000 アイテムから構成される。このデータと観光スポットを対応付けて観光スポットを介したクロスドメイン推薦システムについて検討する。

キーワード 位置情報, 観光情報推薦, クロスドメイン推薦

1. はじめに

2015 年の訪日外国人旅行者数が 1,974 万人に達する^(注1) など観光を楽しむ人々はますます増大してきている。実際に観光地を訪れた人々は、現地での景色を眺めたり、名物料理を食したりするなど、観光の楽しみ方は人それぞれである。中には、観光地に訪れたとき、その観光地が舞台になった小説を読みたい、映画を観てみたいなど、その観光地に関連するアイテムを探すことも観光を楽しむ方法の一つである。

このような背景の下、本研究では観光地を介したクロスドメイン推薦システムの実現を目指している。このシステムでは、ユーザが興味のある観光地を入力すると、その観光地に関連した書籍や映画、楽曲など、ドメインをまたがったアイテムをユーザに提示する。このようなシステムを実現するためには、あらかじめ観光地とアイテムとを関連付けておく必要がある。

本研究では、クラウドソーシングによりアイテムへの位置情報付与を行う。アイテムに位置情報を付与することをジオタギング、位置情報が付与されたアイテムをジオタグアイテムと定義する。これらジオタグアイテムを参照することで、位置情報をベースに観光地とアイテムとを関連付ける。

本研究の貢献は次のとおりである。

- クラウドソーシングにより、小説 9,945 件、映画 9,417 件、楽曲 9,668 件の計 29,030 件のジオタグアイテムを収集した。
- 地図インタフェースを採用した Web システムを開発した。システム上でユーザが地図操作を行うことで、最寄りの観光スポットを検出し、そのスポットに関連するアイテムをアイテムビューに提示する。

2. 関連研究

2.1 クロスドメイン推薦システム

従来の推薦システム [1] は単一のドメインにおいて嗜好データを獲得し、それに基づき当該ドメインにおける推薦を行うという方法が主流であった。例えば、映画推薦システムでは、映画の視聴履歴や評価履歴に基づき対象ユーザの嗜好に関する学習モデルを獲得する。その学習モデルを用いてやはり映画ドメインにおける嗜好を予測する。しかし、単一ドメインにおいて獲得された学習モデルは他のドメインにおいては利用できない。仮に映画ドメインにおいて十分な嗜好データが蓄積されていたとしても、それをそのまま書籍ドメインにおける嗜好の予測に用いることはできない。

そこで、近年クロスドメイン推薦という考え方が注目されてきた [2]。クロスドメイン推薦システムは、複数のドメインにまたがった推薦を実行するシステムである。例えば、映画ドメインで蓄積された嗜好データを基に、書籍ドメインにおける嗜好を予測するというものである。これは異なるドメインにおいてもドメイン間の嗜好には何らかの関連性があるという考えに基づく。

クロスドメイン推薦は、一つのドメイン（ソースドメインとよぶ）において蓄積された評価値データを用いて、他のドメイン（ターゲットドメインとよぶ）における評価値を予測することが目的となる [3]。Cremonesi ら [4] は、クロスドメイン推薦問題を図 1 のように四つの類型に分類している。図 1 は、ユーザ-アイテムの評価値行列を表し、行がユーザを列がアイテムを表す。 D_S がソースドメインの評価値行列、 D_T がターゲットドメインの評価値行列を表す。クロスドメイン推薦では、ソースドメイン D_S の評価値行列を基に、ターゲットドメイン D_T の評価値を予測することが課題となる。

ソースドメイン D_S からターゲットドメイン D_T の評価値を

(注1) : http://www.mlit.go.jp/kankochou/siryou/toukei/in_out.html

予測するための一つの手がかりとして、共通ユーザ集合および共通アイテム集合が用いられる．図 1(b) は、 D_S, D_T において共通ユーザ集合が存在することを表す．つまり、同一のユーザが異なるドメインにおいても評価値を与えていることを意味する．図 1(c) は、 D_S, D_T において共通アイテム集合が存在することを表す．つまり、同一の（あるいは同等とみなせる）アイテムが異なるドメインにおいて存在することを意味する．図 1(a) は、共通ユーザ集合も共通アイテム集合も存在することを表す．このように共通ユーザ集合や共通アイテム集合が存在する場合、 D_S, D_T において評価値行列が重複するため、その重複部分から D_S と D_T の関係性を推定することが可能である．例えば、Sahebi ら [3] は、この重複部分を基に正準相関分析を行い、ソースドメインの評価値をターゲットドメインの評価値への換算式を獲得している．

このように評価値行列の重複部分が存在するような場合には、ソースドメインからターゲットドメインへの嗜好の予測は可能となるが、図 1(d) のように重複部分が存在しない場合にはこのようなアプローチは適用できない．したがって、二つのドメインにおける共通属性を別途抽出するなどのアプローチが必要である．

本研究では、複数のドメインにおける共通属性として位置情報に着目している．クラウドソーシングによりアイテムに位置情報を付与しておく．位置情報を基に観光スポットと各ドメインに属するアイテムとを対応付けておくことで、観光スポットを介したクロスドメイン推薦が可能であると考えられる．

2.2 観光情報推薦システム

観光情報推薦システム [5] は、ユーザの要求に合った観光情報や娯楽施設を提供するシステムである．これまでにも多くの観光情報推薦システムに関する研究が行われている [6]．

観光情報推薦システムは、Web インタフェースとして提供されることが多い．Lee ら [7] は、Web ページに Google Maps を埋め込み、地図上で旅行ルートを表示するインタフェースを提案している．台南市での旅行者向けに、文化遺産や郷土料理など個人の嗜好に合った旅行プランを提示する．EnoSigTur [8] では、ユーザが自身のでもグラフィック情報および嗜好を入力することで、観光に関する情報が推薦される．

さらに、スマートフォンなどのモバイルインタフェースとして観光情報推薦システムを適用しているものもある．Yu ら [9] は、PDA において個人向け旅行プランニングをサポートするための位置情報ベースの推薦システムを提供している．推薦情報は個人の嗜好や位置情報、時間に基づいて生成され、Google Maps 上に提示される．LiveCities [10] は、Android での通知サービスを用いて、ユーザコンテキストに合った情報をプッシュ提供する．

以上のように、観光情報推薦システムは Web インタフェースやモバイルインタフェースとして提供されていることが多い．特に観光情報は地域との関係が強いため、推薦情報を提示するために地図インタフェースを採用しているシステムも多い．本研究においても、地図インタフェースを採用している．

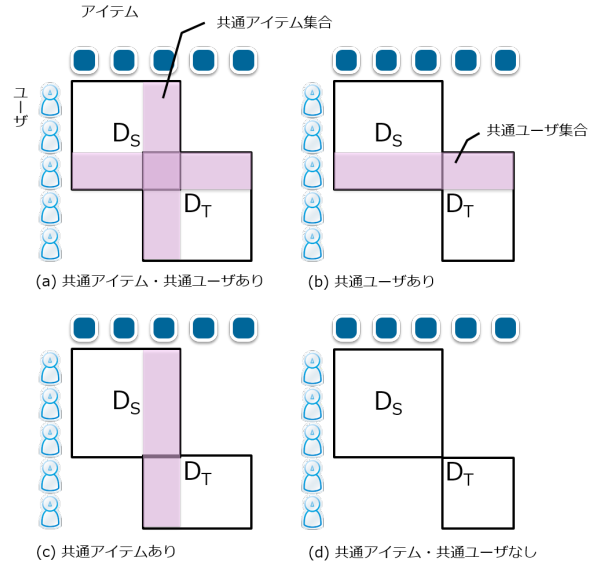


図 1 クロスドメイン推薦問題の類型．二つの異なるドメイン（ソースドメイン D_S およびターゲットドメイン D_T ）において評価値行列が与えられたとき、ドメイン間での共通アイテムの有無、共通ユーザの有無によって四つの類型に分類される．

3. 観光スポットベースクロスドメイン推薦システム

本章では、提案システムである観光スポットベースクロスドメイン推薦システムについて説明する．まず、システムのインタフェースについて説明し、システム構成について説明する．

3.1 インタフェース

図 2 はシステムのインタフェースである．インタフェースは、大きく、地図ビューおよびアイテムビューから構成される．ユーザは地図ビューにおいて地図操作（移動、ズームイン、ズームアウト）を行うことができる．地図操作に応じて、アイテムビューに提示されるアイテムが動的に変化する．

アイテムビューには、表示されている地域に応じて、その地域に関連する本、映画、楽曲が提示される．図 2 の例では、京都に関連するアイテムとして、小説『信長新記 1 本能寺炎上』や映画『大奥』、楽曲『京都の大学生』が提示されている．

3.2 システム構成

図 3 にシステム構成を示す．以下、各処理について説明する．項目の番号は図中の番号に対応する．

(1) Web から観光スポットおよびアイテムに関するデータを収集する．観光スポットデータはスポットテーブルに、各種アイテムに関するデータはそれぞれのアイテムテーブルに格納する．データの収集には各種の WebAPI を用いる．本研究では、観光スポットデータの収集には Foursquare API^(注2) を、書籍データの収集には Google Books APIs^(注3)、映画および楽曲データの収集には iTunes Search API^(注4) をそれぞれ用いた．

(注2) : <https://developer.foursquare.com/>

(注3) : <https://developers.google.com/books/>

(注4) : <https://affiliate.itunes.apple.com/resources/documentation/itunes-store-web-service-search-api/>



図 2 システムのインターフェース。地図ビューおよびアイテムビューから構成される。ユーザが地図ビューにおいて地図操作（移動，ズームイン，ズームアウト）を行うと，それに応じてアイテムビューに提示されるアイテムが動的に変化する。

(2) 収集したアイテムデータについて，関連する地域に基づき位置情報（ジオタグ）を付与する．アイテムとジオタグの関係をアイテム-ジオタグテーブルに格納する．アイテムへの位置情報付与（ジオタギング）は，クラウドソーシングにより行う．クラウドソーシングによるアイテムへのジオタギングについては，4. 章で詳細に説明する．

(3) スポットの位置情報およびアイテムのジオタグに基づき，スポットとアイテムとを関連付ける．関連付けたスポットとアイテムの対を，スポット-アイテムテーブルに格納する．スポットとアイテムとの関連付けは 5. 章で詳細に説明する．

(4) 地図領域の中心に最も近いスポットを入力スポットとし，そのスポットをキーとし，スポット-アイテムテーブルから関連アイテムを取得する．関連アイテムは各ドメインから 3 件ずつ取得する．取得した関連アイテムをアイテムビューに提示する．

4. クラウドソーシングによるアイテムへのジオタギング

小説や映画，楽曲の各ドメインに属するアイテムに対し，その内容に関連する位置情報を付与する．ここで，内容に関連する位置情報の例として，ドメインごとに次のように示す．

小説 小説の舞台となっている地域や小説の中で紹介されている地域，主人公ゆかりの地域など．

映画 映画の舞台となっている地域や映画のモデルとなった地域，映画の中で言及されている地域，主人公ゆかりの地域，映画のロケ地など．ただし，製作国や公開国，監督の出身地など，映画の内容に関係しないものは，関連する地域とはみなさない．

楽曲 この地域が歌詞に現れる，この地域をテーマとして作曲された，この楽曲がテーマソングとして使われたドラマや映画の舞台になっているなど．ただし，楽曲が演奏されたコンサー

ト会場やアーティストの出身地など，楽曲の内容に関係しないものは，関連する地域とはみなさない．

以上のような内容に関連する位置情報をアイテムに付与するためには，そのアイテムに関する十分な知識が必要となる．そこで，本研究では，クラウドソーシングによりアイテムへの位置情報付与を行う．なお本論文では，アイテムに位置情報を付与することをジオタギング，位置情報が付与されたアイテムをジオタグアイテムと定義する．

4.1 タスク設定

図 4 にクラウドソーシングによるアイテムへのジオタギングタスクのイメージを示す．図には記入例も含めている．図の例では，ワーカが小説『信長新記 本能寺炎上』（ISBN:9784198920609）に対し，京都の本能寺辺りの緯度・経度を付与していることを表す．なお，緯度・経度の参照には，外部サイト『Google マップで緯度・経度を求める』^(注5) を利用してもらった．

また，ワーカには，先に挙げたアイテムの内容に関連する位置情報の考え方を示している．さらに，広い地域に関連すると思われる場合は，その地域内の任意の地点を入力すること，複数の地域に関連すると思われる場合は，同一の小説について複数回に分けて回答することについて同タスク画面で説明している．映画および楽曲についても同様にタスクを実施した．つまり，ドメインごとにタスクを実施した．

4.2 RDB への登録

本タスクにより得られたデータを RDB に登録する．ドメインごとに次のアイテム-緯度経度テーブルを用意する．

```
book-latlng(worker_id, book_id, latlng)
movie-latlng(worker_id, movie_id, latlng)
track-latlng(worker_id, music_id, latlng)
```

(注5) : <http://user.numazu-ct.ac.jp/tsato/webmap/sphere/coordinates/advanced.html>

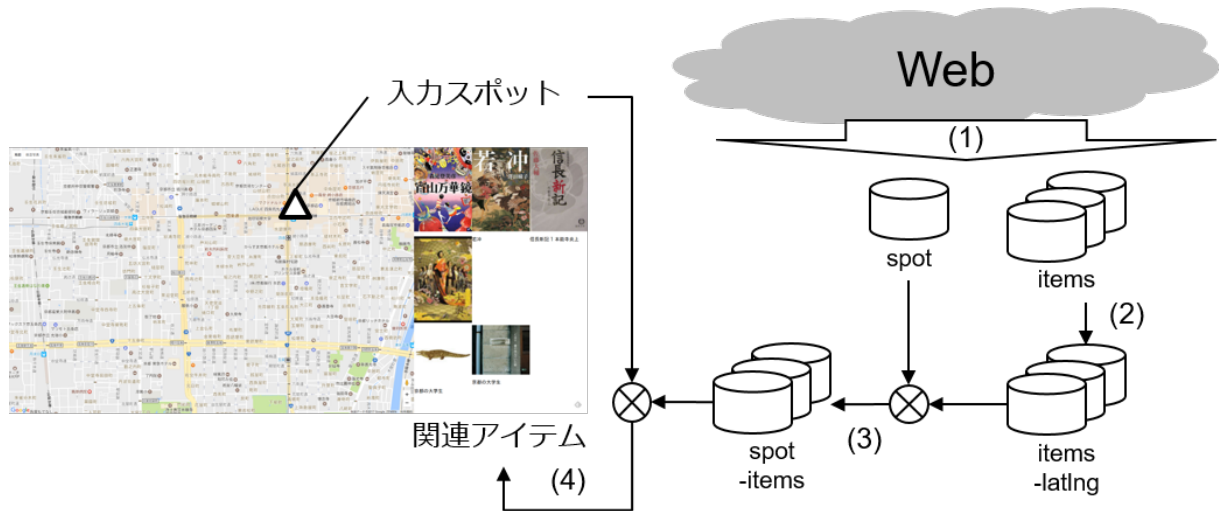


図3 システム構成。(1) 事前に Web から観光スポットおよびアイテムに関するデータを収集しておく。(2) 収集したアイテムについて位置情報を付与 (ジオタギング) する。(3) 観光スポットとジオタグアイテムとを対応付ける。(4) 地図領域の中心に最も近いスポットを入力とし、対応付けられたアイテムを各ドメインから3件ずつ取得し、アイテムビューに提示する。

【質問】
これまでに読んだことがある小説を一つ取り上げ、その小説の内容に関係すると思われる地域の緯度・経度を入力してください。

【回答1】小説のISBN：

【回答2】小説の内容に関連する地域の緯度・経度
緯度：
経度：

図4 クラウドソーシングによるアイテムへのジオタギングタスクのイメージ。

ここで、`worker_id` はタスクを遂行したワーカーの ID を表す。`book_id`, `movie_id`, `track_id` はそれぞれ小説、映画、楽曲のアイテム ID を表す。`latlng` は緯度・経度を表し、実装上はジオメトリ型で扱われる。

また、既に図2の(1)の処理において、各ドメインのアイテムデータが各アイテムテーブル `book`, `movie`, `track` に登録されているものとする。`book-latlng`, `movie-latlng`, `track-latlng` の各アイテム ID は、対応するアイテムテーブルを参照する外部キーである。

4.3 データ分布

本タスクはドメインごとに10,000件依頼した。そのうちアイテムIDに不整合がある、緯度・経度が有効な値でないなどのデータを排除した。得られた有効データ数は、小説9,945件、映画9,417件、楽曲9,668件となった。ユニークデータ数は、小説6,901件、映画5,279件、楽曲6,000件となった。

ジオタグアイテムの分布は図5のとおりとなった。今回はタスク依頼対象が日本人であったため、日本にデータが集中したことは当然である。一方で、ヨーロッパやアメリカ、その他の地域にもジオタグが付与されていることがわかる。

5. スポット関連アイテムの推薦

観光スポット $s_i \in S$ と、各ジオタグアイテム $c_j \in B, M, T$ とを関連付ける。ここで、 B は書籍集合、 M は映画集合、 T は楽曲集合とする。関連付けは、観光スポットの位置情報 $(s_i.x, s_i.y)$ および各ジオタグアイテムの位置情報 $(e_j.x, e_j.y)$ を基に行う。関連付け手法として、(a) 半径ベース手法、(b) k 最近傍法の二つの手法を提案する。以下、各手法について説明する。

5.1 半径ベース手法

観光スポット $(s_i.x, s_i.y)$ を中心とした半径 r の円領域に存在するジオタグアイテム集合を関連付ける。

5.2 k 最近傍法

観光スポット $(s_i.x, s_i.y)$ から、ドメインごとに最近傍の上位 k 件のジオタグアイテム集合を関連付ける。

6. 評価実験

提案システムの有用性を検証するために被験者による評価実験を行った。本章では、被験者評価実験について、データセットおよび比較手法、実験方法について説明したうえで、定量分析および定性分析について述べる。

6.1 データセット

評価実験では、Foursquare API を用いて32,775件の観光スポットを収集した。また、4章で述べた、小説9,945件、映画9,417件、楽曲9,668件のジオタグアイテムを用いた。

6.2 比較手法

比較手法として、5章で説明した下記の二つを用意した。

- (a) 半径ベース手法
- (b) k 最近傍法

6.3 実験方法

被験者は男性10名であり、年齢は21歳から23歳である。

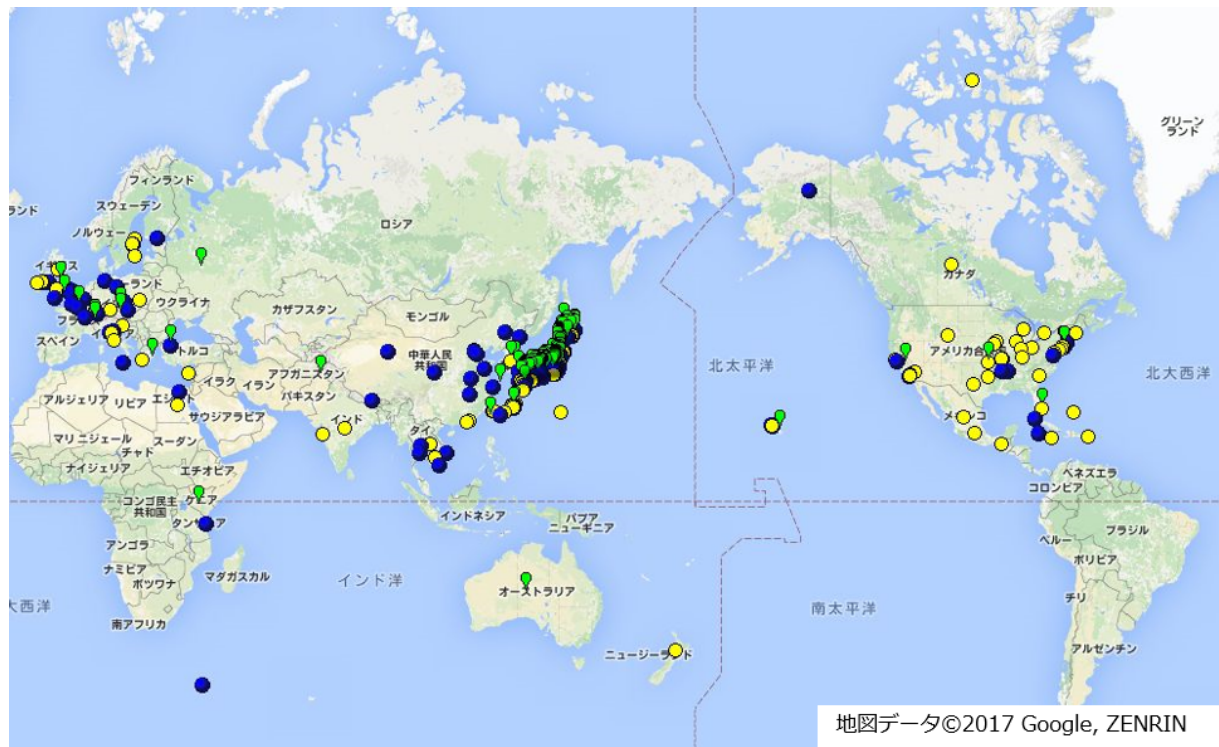


図 5 ジオタグアイテムの分布。

被験者は提案システムに関する予備知識はもたない。被験者にシステムを使用してもらうシチュエーション設定として、下記の 2 種類のタスクを用意した。

T1 「興味のある地域を決めて、その周辺を探索しながら興味のある本や映画、音楽を探してください。」

T2 「面白そうな本や映画、音楽が見つけれそうな地域を探し、そこで興味のある本や映画、音楽を探してください。」

タスク T1 は、興味のある地域から興味あるアイテムを発見しようというシチュエーションである。一方で、タスク T2 は、面白そうなアイテムが見つかりそうな地域にあたりをつけて、地域やアイテムを探索しようというシチュエーションである。

被験者には上記の各タスクを提示し、下記手順にしたがってシステムを使用してもらった。

- (1) 被験者は指定されたシステムを用い指定されたタスクを遂行する（時間は無制限）。
- (2) 被験者はタスクを遂行している間、アイテムビューに提示されたアイテムのうち興味をもったアイテムを記録しておく（記録するアイテムの上限数は無制限）。
- (3) 被験者が満足したか飽きた場合、タスクを終了する。
- (4) 被験者は記録した各アイテムに対して表 1 に示した質問に回答する。
- (5) 他の残りのシステムおよびタスクについても同様の手順を行う。

なお、タスク、システムともに順所効果を相殺するため、タスク、システムの利用順序は被験者により変更している。

ここで、表 1 の質問項目について、Q1 に対しては {YES, NO} の二択で回答してもらう。Q2 から Q5 に対しては {5: 強くそう思う, 4: そう思う, 3: どちらともいえない, 2: そうは

思わない, 1: 全くそう思わない} の 5 段階評価で回答してもらった。

6.4 定量分析

図 6 にタスクごとの結果を示す。図 6 には、表 1 の Q2 から Q5 に対応する評価値を全被験者で平均した値を示している。全体的には、手法間やタスク間では大きな差異はみられなかった。

いずれの手法、いずれのタスクにおいても、共に Q2: 提示前興味よりも Q3: 提示後興味の方が評価値が向上した。この結果より、システムにより観光地に関連したアイテムを提示することで、興味の促進につながるということが示唆される。提案システムの地図インタフェースを介して、アイテムを探索することで、本来、想定していなかったアイテムを発見する可能性が高まり、特に、興味のある観光スポットに関連するという点で、提示されるアイテムにも興味をもてたといえる。

また、いずれの手法、いずれのタスクにおいても、Q5: 逆自力発見性については、平均評価値 4 を超えた。この結果より、観光地に関連したアイテムを提示することで、ユーザにとって普段目に触れないようなアイテムを見つけ出すことができるといことが示唆される。

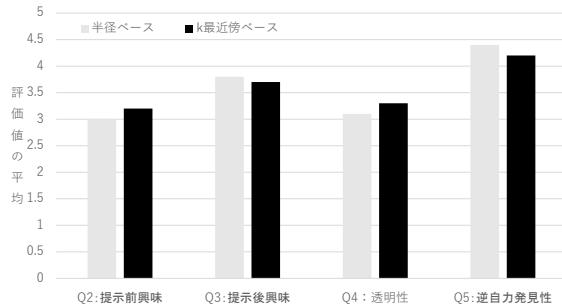
7. 結 論

本研究では、観光地を介したクロスドメイン推薦システムを提案した。システムでは、ユーザが地図操作を行うことで、近傍の観光スポットに関連するアイテムが提示される。このシステムを実現するために、事前にクラウドソーシングによりアイテムへの位置情報付与を行い、収集されたジオタグアイテムの位置情報をベースに観光地とアイテムとの関連付けを行った。

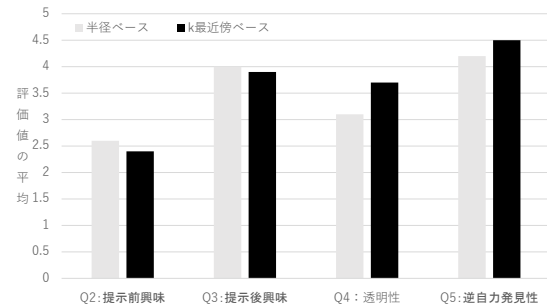
被験者評価を行った結果、提案システムにより提示された各

表 1 発見したアイテムに関する質問項目

質問番号	質問
Q1 (既知)	私はこの { 本, 映画, 音楽 } を知っていた.
Q2 (提示前興味)	システムから提示される前から, 私はこの { 本, 映画, 音楽 } に興味があった.
Q3 (提示後興味)	システムから提示されて初めて, 私はこの { 本, 映画, 音楽 } を読んでみたくなった, 映画を観てみたくなった, 音楽を聴いてみたくなった }.
Q4 (透明性)	私はこの { 本, 映画, 音楽 } がなぜシステムから提示されたかを理解できた. }
Q5 (逆自力発見性)	このシステムを使わなければ, 私自身, 自力ではこの { 本, 映画, 音楽 } に会うことはなかったと思う.



(a) T1



(b) T2

図 6 タスクごとの評価結果

アイテムについて, 提示前興味よりも提示後興味の方が評価値が向上したことを確認した. この結果より, システムにより観光地に関連したアイテムを提示することで, 興味の促進につながるということを示唆した. また, 逆自力発見性についても評価値が高ったことを確認した. この結果より, 観光地に関連したアイテムを提示することで, ユーザにとって普段目に触れないようなアイテムを見つけ出すことができるということを示唆した.

今後は, 観光地とアイテムとの関連付け方法の洗練, ユーザインタフェースの改良等を行い, 観光地を介したクロスドメイン推薦システムの改良していく.

謝 辞

本研究は JSPS 科学研究費 16HO5932 の助成を受けたものです. ここに記して謝意を表します.

文 献

- [1] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor. *Recommender Systems Handbook*. Springer, 2011.
- [2] Ignacio Fernández-tobías, Iván Cantador, Marius Kaminskis, and Francesco Ricci. Cross-domain recommender systems : A survey of the State of the Art. In *Proceedings of the 2nd Spanish Conference on Information Retrieval*, pp. 187–198, 2012.
- [3] Shaghayegh Sahebi and Peter Brusilovsky. It Takes Two to Tango : an Exploration of Domain Pairs for Cross-Domain Collaborative Filtering. In *RecSys 2015: Proceedings of the 9th ACM conference on Recommender systems*, pp. 131–138, 2015.
- [4] P. Cremonesi, A. Tripodi, and Turrin. R. Cross-domain recommender systems. In *ICDMW2011: IEEE 11th International Conference on Data Mining Workshops*, pp. 496–503, 2011.
- [5] Francesco Ricci. Travel Recommender Systems. *IEEE Intelligent Systems*, Vol. 4, No. December, pp. 55–57, 2002.
- [6] Joan Borràs, Antonio Moreno, and Aida Valls. Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, Vol. 41, pp. 7370–7389, jun 2014.

- [7] Chang Shing Lee, Young Chung Chang, and Mei Hui Wang. Ontological recommendation multi-agent for Tainan City travel. *Expert Systems with Applications*, Vol. 36, No. 3 PART 2, pp. 6740–6753, 2009.
- [8] Joan Borràs, Joan de la Flor, Yolanda Pérez, Antonio Moreno, Aida Valls, David Isern, Alicia Orellana, Antonio Russo, and Salvador Anton-Clavé. SigTur/E-Destination: A System for the Management of Complex Tourist Regions. In R. Law, M. Fuchs, and F. Ricci, editors, *International Conference on Information and Communication Technologies in Tourism, ENTER 2011*, Vol. 3, pp. 39–50. Springer, 2011.
- [9] Chien Chih Yu and Hsiao Ping Chang. Personalized location-based recommendation services for tour planning in mobile tourism applications. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 5692 LNCS, pp. 38–49, 2009.
- [10] D Martin, a Alzua, and C Lamsfus. A contextual geofencing mobile tourism service. In R. Law, M. Fuchs, and F. Ricci, editors, *Information and communication technologies in tourism 2011*, pp. 191–202. Springer-Verlag, 2011.