

# 仮想空間における滞在時間を用いたコンテンツベース推薦手法の提案

## Content-based Recommendation Method using Visiting Time for Second Life Virtual Worlds

森 和貴†      山竹 浩史†      レハイ バン†      ターウオンマツ ラック†      原田 智広†  
Kazuki Mori      Hirofumi Yamatake      Bang Hi Le      Ruck Thawonmas      Tomohiro Harada

### 1. はじめに

近年情報技術の急速な発展に伴い、個人や団体が容易にかつ低コストで情報を発信することが可能になり、多くの情報を用意に獲得可能になっている。しかし、情報が膨大になるにつれ、類似した情報が大量にありどれが本当に自分の望む情報か判別することができないといったような「情報過多」の状態に陥り、本当に必要な情報を見つけ出すことは困難になる。これを打破するために、膨大な情報から必要な情報のみを提示する推薦システムが提案されている[1]

また近年では、メタバースと呼ばれるインターネット上に構築された多人数同時参加型の3次元仮想空間が注目を集めている。メタバースは他のオンラインゲームとは異なり行動シナリオなどユーザの行動を制約するものではなく、ユーザは自由にメタバース内を行動できる。このような特徴から近年では、娯楽としてだけでなく企業でのマーケティング利用や教育への導入などでメタバースに注目が集まっている[2][3]。

メタバースの代表的なサービスとして **Second Life** (以下, **SL**) がある。SL 内は多数のエリア群から構成され、ユーザはエリアを自由に移動することができる。しかし、SL ではあまりにも構成するエリア数が多いためユーザはどのエリアが本当に興味あるエリアなのかを判別することができないのが現状である。そこで本研究では SL におけるエリア推薦手法として、ユーザのエリアごとの滞在時間に着目した方法を提案する。

### 2. Second Life

SL とは 2003 年より Linden Lab 社によって運営されている 3DCG で構成された仮想空間である。SL は 400 万人以上(2014/12)のユーザが登録しており世界で最大規模のメタバースである。SL には Linden Script Language(以下 LSL)という独自言語が実装されておりユーザはこれを用いて自由にオブジェクトの生成、制御ができる。SL 内に存在するオブジェクトはほとんど LSL にとって作られており、オブジェクトの所有権は製作者に帰属する。また、独自の通貨による LSL で生成されたオブジェクトの売買やエリアと呼ばれる土地のレンタルも可能である。これらの要素からユーザの自由度が高い仮想空間であるとして世界中から注目を集めている。

しかし、エリアの数は膨大で公式ホームページである「Destination Guide」だけでも 2347 エリア(2015/3/9)が紹介されている。このためユーザが自身の好むエリアを探すのは困難である。

### 3. 推薦システム

推薦システムとは多数の情報の中から本当にユーザが求めている情報を提供するというものである。推薦シス

テムには大きく分けて協調フィルタリングとコンテンツベースフィルタリングの2種類の方法がある。協調フィルタリングとは多くのユーザの中から趣向の近いユーザを見つけ出し、そのユーザの趣向データをもとに推薦する手法である。コンテンツベースフィルタリングはアイテム間の類似度をあらかじめ計算したプロファイルを作成しそれをもとに推薦する手法である。

協調フィルタリングとコンテンツベースフィルタリングの比較をすると、協調フィルタリングは他人のデータをもとに推薦を行うため、自分が知らない特徴を持つアイテムの推薦が可能である。また協調フィルタリングはアイテムの特徴データを利用してないため新しいアイテムを追加するたびにアイテム特徴を計算しなおす必要性がない。これに対しコンテンツベースフィルタリングは、あらかじめ作成しておいたプロファイルを基に推薦をするので、他人のデータを使う協調フィルタリングと異なり利用者数が少なくても推薦することができる。またあらかじめプロファイルでアイテム類似度を計算しているので、新しいアイテム、人気のないアイテムなどあまりユーザから評価されてないアイテムも推薦できる。

メタバースでは常に新しいアイテムが追加され、全てのアイテムがユーザの評価を受けているとは限らない。そこで本研究ではコンテンツベースフィルタリングを用いて推薦を行った。

また推薦に用いるユーザのアイテム (SL ではエリア) に対する興味度として明示的評価と暗黙的評価の2種類が存在する。明示的評価とは、アンケートなどによって得られるユーザの直接的な評価を使用する手法である。暗黙的評価とは、ユーザの行動履歴などからユーザの興味度を推察する手法である。明示的評価はユーザの嗜好を直接反映できるため評価としては正確である。一方、暗黙的評価は、ユーザに負荷がかからないため多くのデータを集めることが可能である。

本研究では、エリアに対するユーザの興味度の暗黙的評価として滞在時間を利用する。滞在時間が長いエリアほどそのエリアに興味を持っていると考えられており[4]、これを暗黙的評価として利用できると考えたからである。

### 4. 提案手法

コンテンツベースフィルタリングでエリア推薦を行うために、エリア間の類似度を計算する必要がある。類似度を計算するために、Destination Guide から得られる各エリアの説明文である「エリアプロフィール」を用いる「エリアプロフィール」から tf-idf を計算し、各エリアの tf-idf 値に基づいてエリア間の類似度を、コサイン類似度を用いて求める。

コンテンツベースフィルタリングではこれらの要素を用いて、式(1)に示す Sarwar の方法[5]を用いて未訪問エリアの予測評価値を計算し、その値が大きいエリアをユーザに対して推薦する。ただし、 $s_{i,j}$  は訪問済みエリア  $j$  と

†立命館大学, Ritsumeikan University

本研究の一部は文部科学省科学研究費補助金(基盤研究(C), 課題番号: 26330421)による助成による援助を得て実施された。

未訪問エリア  $i$  間の類似度,  $r_{u,j}$  はユーザ  $u$  の訪問済みエリア  $j$  に対する評価値,  $\hat{r}_{u,i}$  はユーザ  $u$  の未訪問エリア  $i$  に対する予測評価値を示す.

$$\hat{r}_{u,i} = \frac{\sum_j s_{i,j} \cdot r_{u,j}}{\sum_j s_{i,j}} \quad (1)$$

ここで, 本研究ではユーザのエリアに対する評価値としてエリアの滞在時間を用いる. しかし, SL では“エリア”ごとの面積が異なるため滞在時間を評価値として用いる場合, 広いエリアほど滞在時間が大きくなるため適切な評価ができない. そこで, エリア面積の差による滞在時間の差を考慮し, 式(1)に新しい項を追加した, 式(2)を新しい予測評価値の式として導入する

$$\hat{r}_{u,i} = \frac{\sum_j s_{i,j} \cdot r_{u,j} \cdot \frac{size_{max}}{size_{min}}}{\sum_j s_{i,j}} \quad (2)$$

$size_{max}$  は推薦候補にあるエリアの中で最も大きいエリアの面積,  $size_j$  は訪問済みエリア  $j$  の面積を表す. 最大面積との比を計算し, それを滞在時間にかけることで, 広いエリアの滞在時間は変えず, 狭いエリアの滞在時間を引き伸ばし, 広いエリアに合わせ正規化する.

## 5. 実験結果

滞在時間がユーザ嗜好の暗黙的趣向として有用性があるかを調べるために, 20人のユーザに対してSLを用いたエリア推薦実験を行った. 被験者は著者の所属する大学内の学生から募集し年齢層は21歳から24歳である. いずれの被験者もSLをプレイするのは初めてである.

### 5.1 比較手法

Destination Guideより取得した各エリアの「いいね」を取得し, これの高い順にエリアを推薦する. 「いいね」とはフェイスブック株式会社 (Facebook, inc) が提供するソーシャルネットワークサービスで用いられる好感を示すコミュニケーションツールで, エリアに対して好感が持てたときに「いいね」と評価する. この様な特性から「いいね」はユーザの明示的であると捉えられる.

### 5.2 被推薦エリア群

本実験に用いる被推薦エリア群は Destination Guide の各カテゴリのうち最も「いいね」の数が多い40のエリアとする.

### 5.3 実験内容

以下に実験の流れを示す.

Step1被験者は用意された被推薦エリア群の中から1つランダムに訪問する.

Step2被験者は満足するまで先ほど選ばれたエリアを観光する.

Step3被験者は訪問したエリアに対して1から5までの5段階で評価する.

Step4Step1からStep3までを3回繰り返す

Step5以下の(a)と(b)を交互に2回ずつ行う. 最初にどちらを行うかは毎回ランダムに決定する.

#### (a) 提案手法を用いた推薦

Step(a)-1. (2)式を用いて計算された予測評価値が最も高いエリアを推薦する.

Step(a)-2. 被験者は満足するまで選ばれたエリアを

観光する.

Step(a)-3. 被験者は訪問したエリアに対して1から5までの5段階で評価する.

#### (b) 比較手法を用いた推薦

Step(b)-1. まだ訪れていないエリアの内最も“いいね”の数が多いエリアを推薦する.

Step(b)-2. 被験者が満足するまで選ばれたエリアを観光する.

Step(b)-3. 被験者に訪問したエリアに対して1から5までの5段階で評価する.

## 5.4 実験結果

表1: 各種法における合計評価値の平均

| 比較手法の平均合計評価値 | 提案手法の平均合計評価値 |
|--------------|--------------|
| 6.1          | 7.3          |

表1は上記の実験(a), (b)各手法で訪れた, それぞれ2エリアずつの合計評価値のユーザ平均を表している. 表1より提案手法のほうが比較手法に対して平均して高い評価を得られたことが分かる. また各ユーザの滞在時間と評価値の相関を計算したところ, 平均相関係数  $r = 0.373$  となりユーザのエリア滞在時間とエリア評価値の間にはある程度の相関がみられた.

## 6. 考察

「結果の各ユーザにおける手法ごとの評価合計値に差がない」という帰無仮説を立てウィルコクソンの符号付順位検定を行った結果,  $p=0.017$  となり2つの手法の差は優位であるという結果が得られた. このことからSLエリア推薦において, 滞在時間がユーザの暗黙的評価として利用できることが示された.

## 7. まとめ

本稿では, メタバース内のエリア推薦に向け, 滞在時間をユーザの暗黙的評価として利用し, Sarwarの方法の改良式を提案し, SLを用いたエリア推薦実験を行った. 実験の結果, 滞在時間が暗黙的評価として利用できること明らかになった. 今回の実験ではエリア類似度にエリアプロフィールの tf-idf 値を用いたコサイン類似度を用いたが, 今後は土地のカテゴリや地形特徴を考慮した類似度を算出することにより, より精度の高い推薦を行うことを目指す. また今回は「いいね」を用いた手法とのみ比較したが, 今後は他の手法との比較も行い優位性を確かめていきたい.

### 参考文献

- [1] 神嶋 敏弘, 推薦システムのアルゴリズム, 人工知能学会, 誌, vol.22-no.6-2007, vol.23-no.1-2008, vol.23-no.2-2008.(2008)
- [2] 木下まどか, and 森本祥一. “メタバースにおけるビジネスモデルとその効果に関する考察.” 経営情報学会全国研究発表大会要旨集 2011.0 (2011): 740-740.
- [3] 小川真里江, and 新井正一. “メタバースを活用した授業.” コンピュータ & エデュケーション 28.0 (2010): 45-48.
- [4] Bohnert, Fabian, et al. "Using interest and transition models to predict visitor locations in museums." AI Commun. 21.2-3 (2008): 195-202.
- [5] Sarwar, Badrul, et al. "Item-based collaborative filtering recommendation algorithms." Proceedings of the 10th international conference on World Wide Web. ACM, 2001.