

外出プラン作成の為の嗜好情報収集システムの研究

A Study about Collecting Preference Information System for Creating Outing Plan

山下 和也† Kazuya Yamashita 岩本 健嗣‡ Takeshi Iwamoto 松本 三千人‡ Michito Matsumoto

1. はじめに

今日、旅行者を対象とした外出時における様々なプランニングシステムが存在する[1]。これらは、旅行やドライブをする為の外出プランを事前に作成し、サポートをするシステムである。既存のプランニングシステムは、ユーザの好みに合った店舗候補や立ち寄り先を提示するなど、嗜好情報を利用することを目的としている。しかし、嗜好情報を把握する為に、ユーザに選択項目を選ばせるといった手間があり、また、選択項目数によっては、十分に嗜好を把握できないといった問題がある。そこで、本研究では、日常生活での移動履歴を、スマートフォンに搭載された GPS を用いて取得する。この日常の移動履歴から嗜好情報を収集する。そして、収集した嗜好情報をプランニングシステムに用いることで、問題点を解決する。

2. 嗜好情報

既存の研究では、嗜好情報をユーザの音楽情報検索履歴や、事前に抽出した個人の好みに関する情報を基に定義している[2][3]。本研究においては嗜好情報を、ユーザが普段好んで利用する施設情報、好みの料理や飲食店情報とする。これらの情報を収集し、利用することで、プランニングシステムにユーザの嗜好を反映させることができる。例えば、普段利用する施設情報からは、施設に対する移動履歴頻度を考慮することで、プランニングの際にユーザに適切な観光スポットや名所などを推薦可能となる。また、好みの料理や飲食店情報を把握することで、よりユーザの嗜好を反映させることが可能となる。

3. システム概要

本研究では、スマートフォン上で GPS 情報を収集することで移動履歴情報の取得を行う。しかし、GPS 情報は屋内において程度の大きな誤差が生じる可能性がある為、主に滞在する施設を取得する本研究では十分な結果が得られないと考えられる。そこで、この問題を解決する為に本研究のシステムでは、収集した GPS 情報に対してクラスタリングを行い、クラスタに対して属性情報を付加することで、滞在した施設の属性情報を嗜好情報として取得する。本手法の概要を図 1 に示す。まず、クラスタリングによって、GPS 情報から、滞在した施設を抽出する。次に、抽出した施設の位置を、実際に存在する施設情報と照合することで、施設の属性情報を嗜好情報として収集する。そして、収集した嗜好情報をデータベースへ格納することで、プランニングシステムにユーザの嗜好情報を反映させることが可能となる。このシステムに

より、ユーザの嗜好情報を詳細に把握することができる。

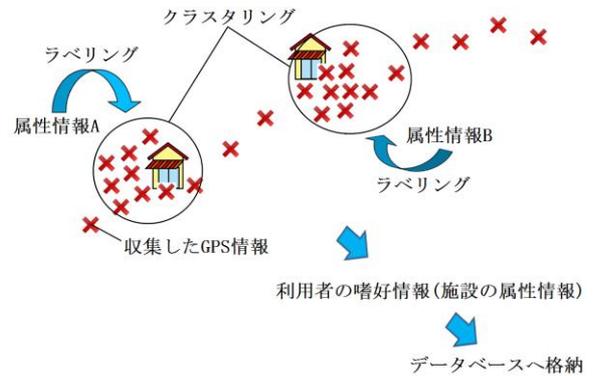


図 1. システム概要図

4. 設計と実装

本章では、クラスタリングを用いた嗜好情報収集手法について述べる。まず、スマートフォンを用いて GPS 情報と時間情報を 30 秒間隔で収集する。次に、収集した GPS 情報とユークリッド距離を用いてクラスタリングを行う。また、クラスタリングの際の類似度の計算には、本研究ではスマートフォンによるシステム構成を考えている為、計算コストが小さく、高い類似度が得られる最遠距離法を用いる。しかし、この手法だけでは、同じ場所に時間間隔を置いて何度も訪れた場合に、実際には訪れていないが、システムによって訪れたと判断してしまい、正しく移動履歴情報を取得できない場合がある。例えば、同じ交差点に何度も留まった場合、短い滞在時間が蓄積し、一定の滞在時間を超すことでシステムが実際に訪れた施設と判断する為である。そこで、本手法ではこの問題点を解消する為に、収集した GPS 情報に関する時系列を考慮するアルゴリズムを設計し、クラスタリングに取り入れることで問題点を解消した。そして、クラスタリング結果によって取得した移動履歴情報の GPS 情報に対して、本手法では Foursquare の API を利用することで、ラベリングを行う。ラベリングでは、取得した移動履歴情報に施設の属性情報をラベリングすることで、移動履歴情報からユーザの嗜好情報を収集し、データベースへの格納を行う。

5. 評価

クラスタリングを用いた嗜好情報収集手法によって、正しく移動履歴情報を取得可能か確認する必要がある。そこで本章では、クラスタリングによって判断したクラスタ数と、訪れた施設数を正しく取得可能か評価を行った後、時系列を考慮したクラスタリングの評価を行う。クラスタリングに用いる GPS 情報には、ロガーによって 30 秒間隔で収集した 7 日分の GPS 情報を用いる。表 1 に今回取得したデータの概要を示す。

† 富山県立大学大学院情報システム工学専攻, TPU

‡ 富山県立大学情報システム工学科, TPU

表 1. データ概要

1 日の平均移動距離	29.6km
全体の総移動距離	207.1km
1 日の平均滞在場所数	4 カ所
全体の総滞在場所数	28 カ所

クラスタリング手法では、1つのクラスタにおける滞在時間と、同一クラスタと判断する範囲をパラメータとして用いて評価を行う。滞在時間については、1つのクラスタの想定滞在時間を 3, 5, 10, 15, 20 分とする。また、同一クラスタと判断する範囲については、地図上のクラスタ間の距離に対して、各クラスタ間の距離を 1m から 100m まで 1m 間隔で変化させクラスタリングを行った。その結果、滞在時間を 3, 5, 10 分とそれぞれ設定した場合には、特定のクラスタ間の距離において、訪れた施設数に対する正しいクラスタリング結果を得ることができたが、滞在時間を 15, 20 分にそれぞれ設定した場合には、正しい結果を得ることができなかった。これは、今回のデータでは、1つの施設に長時間滞在することが少ない為だと考えられる。

次に、この実験結果を踏まえて、各滞在時間のパラメータを 3, 5, 10 分とし、同一クラスタと判断する時間をパラメータに加えて、時系列を考慮したクラスタリングを行う。同一クラスタと判断する時間とは、設定した時間を超えて同じ場所に何度も訪れた際に、1つのクラスタを複数のクラスタと判断するしきい値である。この時間をそれぞれ 1, 2 時間に設定してクラスタリングを行った。各滞在時間における、クラスタ間の距離と正答率の関係を、同一クラスタと判断する時間がそれぞれ 1 時間の場合を図 2-1 に、2 時間の場合を図 2-2 に示す。正答率とは、各パラメータにおいて、実際に訪れた施設数とシステムが判断したクラスタ数が正しい場合を正答とし、それを割合で表したものである。図 2-1 と図 2-2 より、クラスタ間の距離が遠い場合に、正答率が高くなるのが分かる。これは、クラスタ間の距離を短くすることで、クラスタを余分に分割している為と考えられる。さらに、図 2-1 と図 2-2 を比較してみると、クラスタ間の距離が 60~100m の範囲において正答率に大きな違いが生じ、図 2-2 の方が高い正答率を示している。同一クラスタと判断する時間を短く設定することで、設定した時間を超えて 1つの施設に滞在した場合に、余分にクラスタを分割してしまう場合があった為と考えられる。今回のクラスタリング結果では、想定滞在時間を 5 分と設定し、クラスタ間の距離が 88~98m の範囲かつ、同一クラスタと判断する時間を 2 時間に設定した際に、最も適切なクラスタリング結果を得ることができた。

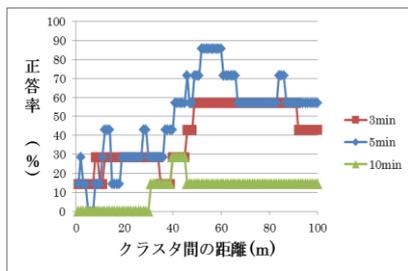


図 2-1. 距離と正答率, 及び滞在時間の関係 1

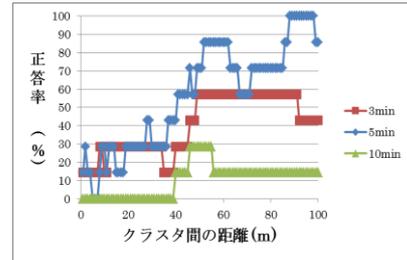


図 2-2. 距離と正答率, 及び滞在時間の関係 2

また、実在する施設の GPS 情報と、クラスタリングによって判断したクラスタの GPS 情報との誤差について解析した。判断したクラスタの GPS 情報には、クラスタ内に含まれる総 GPS 情報の平均値を使用する。図 3 に滞在時間と誤差の関係を示す。図 3 より、滞在時間に比例して誤差の範囲が減少する傾向がみられる。これは、短い滞在時間の場合、誤差の大きい GPS 情報を含んだ際に、判断したクラスタの GPS 情報の平均値が大きく変化すると考えられる。

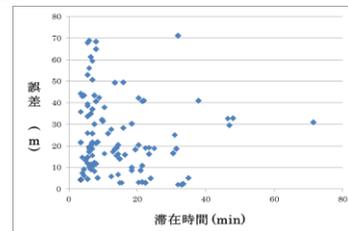


図 3. 滞在時間と誤差の関係

6. まとめ及び今後の課題

本研究では、嗜好情報を把握する為に必要な選択項目数や手間を削減する為に、日常生活から収集した GPS 情報をクラスタリング及びラベリングを行うことでユーザの嗜好情報の収集を行った。クラスタリングを用いた嗜好情報収集手法では、一定の条件下において適切な移動履歴情報を取得することができ、取得した移動履歴情報に施設の属性情報をラベリングすることで、ユーザの嗜好情報を収集することができる。また、施設の GPS 情報に対する誤差の傾向を把握することで、より正確なラベリングを行うことができると考えられる。

今後の課題として、評価に用いた GPS 情報を増やすことで、より詳細なクラスタリング結果を得る為に必要なパラメータ条件の解析、判断したクラスタの GPS 情報と実際の施設の GPS 情報に対する誤差を考慮してのラベリングが挙げられる。また、今回は外出プラン作成の為に必要な嗜好情報を収集するシステムに留まっている為、収集した嗜好情報を反映させるプランニングシステムを実装する。

参考文献

- [1] Panasonic 旅ナビポータブル SD トラベルナビゲーション <http://panasonic.jp/car/tabinavi/products/SG500/index.html>
- [2] 石先広海, 帆足啓一郎, 菅谷史昭, 甲藤二郎, “ユーザ嗜好に基づく音楽情報検索システムにおける学習データ抽出手法”, 情報処理学会研究報告, 2006-MUS-64(13), pp.73-78, Feb.2006.
- [3] 矢野宏実, 目良和也, 相沢輝昭, “嗜好を考慮した評判情報検索手法”, 電子情報通信学会, NLC2004-38, pp.13-18, Nov.2004.