

ベイジアンネットワークを利用した滞在店舗推定

坂東 拓也 †

佐藤 晴彦 †

小山 聡 †

栗原 正仁 †

† 北海道大学 大学院情報科学研究科

1 研究背景

近年、GPS 受信機を搭載した携帯電話やスマートフォンの普及によって、位置情報から推定されるユーザーの状況に合わせた様々なサービスが提供されている [1] [2]。しかし、ユーザーの習慣的な行動以外の不確実な行動によって生じる状況には完全に対応することは難しい。また、GPS は上空にある複数個の衛星との通信を元に位置情報を算出するという性質上、屋内や電波状況の悪い場所では利用することが出来ないため、その際のユーザーの状況は取得できないという問題点がある。

そこで、ユーザーの不確実な行動や GPS の測位不良が起こった際に、ユーザーが移動しているのか、それとも屋内にいるのか、加えて屋内にいるならばどのようなところにいるのか、というユーザーの状況を何らかの形で取得できれば、よりユーザーの状況に合致したサービスの提供に役立つと考えられる。

このような推定を行うためには不確実性や情報の欠落を補完することが必要である。そのためアプローチとして、事象を確率を用いてモデル化し、定量的に扱う方法がある。本研究では確率的手法の中で、複雑な因果関係の中での不確実性推測、データマイニング、知能情報システムなどに対して有効な手段であるといえるベイジアンネットワークに着目した。

2 研究目的

本研究では、ユーザーの位置情報に加えて、ユーザーの属性情報や Web の閲覧履歴から、ベイジアンネットワークを利用し、ユーザーの滞在状況及び、滞在店舗を確率的に推定することを目指す。また Android 端末を用いて、提案手法を実装したシステムの作成を目指す。

3 提案手法

本章では、ユーザーの滞在店舗を推定するアルゴリズムを提案する。

Inferring visiting shops using a Bayesian network

†Takuya Bando, Masahito Kurihara, Satoshi Oyama, and Haruhiko Sato

†E-mail: bando@complex.ist.hokudai.ac.jp,

kurihara@ist.hokudai.ac.jp,

oyama@ist.hokudai.ac.jp,

haru@complex.ist.hokudai.ac.jp

†Graduate School of Information Science and Technology, Hokkaido University

3.1 アルゴリズム

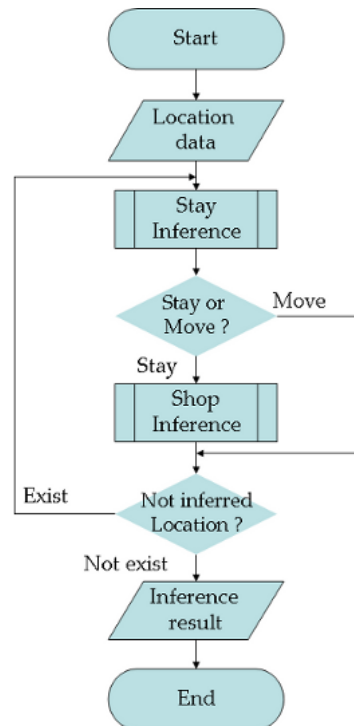


図 1: アルゴリズム

アルゴリズムのフローチャートを図.1 に示す。初めに連続した位置情報系列を入力としてシステムに与える。ここで位置情報とは測位時刻、緯度、経度を情報として持っていることとする。次に与えられた各位置における滞在状況を判定する。更に各位置において滞在していると判定された場合に、滞在している店舗を推定する。これらの操作を未判定の位置情報が存在しなくなるまで繰り返し、最後に推定結果を出力する。各処理の詳細について説明する。

3.2 滞在状況判定

これはユーザーの各位置における滞在状況を推定する処理である。ここで滞在状況とは移動中、滞在中の二つの状況をさす。滞在状況は二つの方法を用いて判定する。一つ目は、位置情報間の距離を用いた判定である。与えられた地点と一つ前の地点の間のユークリッド距離が、閾値よりも小さい場合に、滞在していると判定する (図.2)。

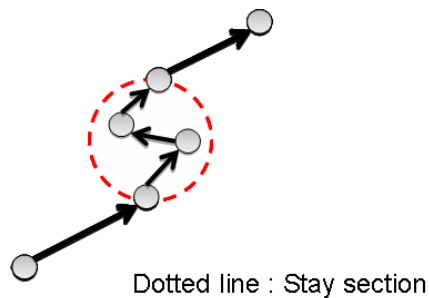


図 2: 距離による判定

それぞれ、与えられた位置を X 、一つ前の位置を X' 、二点間のユークリッド距離を $D(X, X')$ 、閾値を θ とすると、以下を満たす場合に滞在となる。

$$D(X, X') < \theta \quad (1)$$

閾値は GPS の測位間隔と徒歩の平均速度 4km/h から算出する。

二つ目は位置情報の欠落を用いた判定である。これは、GPS の信号が一定時間以上連続で欠落した場合に滞在とみなす処理である (図.3)。一般的に屋内では GPS

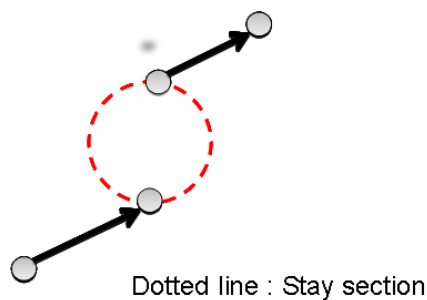


図 3: 欠落による判定

の信号を取得することは難しい場合が多く、情報が欠落しやすい。そのため連続した位置情報の欠落は屋内滞在の可能性が高いと考えられる。

3.3 滞在店舗判定

これはユーザが滞在している店舗を推定する処理である。推定にはベイジアンネットワークを用いる。用いるネットワークの構造を図.4 に示す。

それぞれの確率変数について触れる。初めに滞在店舗とは、ユーザが店舗に滞在している確率を指し、すべての候補となる店舗を含む多値変数である。事前分布は、統計データから得られる各店舗の人気とユーザの Web 閲覧の状況を基に作成する。次にユーザ属性とは、ユーザの年齢、性別を指す。条件付確率は統計データから作成する。最後に座標とは与えられた位置情報と実際の各店舗の位置情報の間の距離を指す。条件付

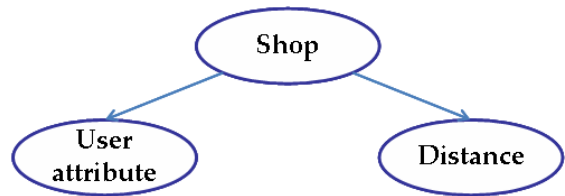


図 4: 提案ネットワーク

確率は、ユーザの位置を中心に正規分布を離散値に変換したものを与える。つまり、近い位置にある店舗に高い確率を与えるものとする。正規分布のパラメータは実験環境や実験端末の性能に基づいて決定する。

4 実験

実データを用いた数値実験を行う。本稿で提案した手法とベイジアンネットワークを用いずに、位置情報と店舗情報のみを用いて、滞在状況と滞在店舗を推定する手法を用意し、比較実験を行う事で提案手法の優位性を証明する。実データを収集する際に、網羅的に実験を行うことは構想システムの中の店舗情報データベースの規模を考えると現実的ではないため、推定対象のジャンルや実験エリアを限定する事で対処する。

5 今後の展望

本稿では位置情報とユーザの情報から、確率的にユーザの滞在状況及び、滞在店舗を推定するシステムを提案した。今後の展望としては構想システムを実際に実装する必要がある。加えて実データの収集と構想システムの優位性を確認するための実験を行う。

参考文献

- [1] 生田目宏昭, 神戸英利, 三井浩康, 小泉寿男. 歩行履歴情報を基にした歩行者ナビゲーションシステムの構築 (セッション 1: 履歴, 状況・行動推定 i). 情報処理学会研究報告. データベース・システム研究会報告, Vol. 2008, No. 7, pp. 13-18, 2008-01-24.
- [2] 宮崎雄一朗, 山田直治, 住谷哲夫. ユーザの行動に合わせたサービス実現のための行動推定技術の開発. NTT DoCoMo テクニカル・ジャーナル, Vol. 17, No. 3, pp. 55-61, 2009-10.