

行動予測に基づく未来のイベントの宣言的かつ連続的な 問い合わせ手法の提案

河村 光則¹ 齊藤 裕樹¹

概要：GPSをはじめ、無線強度や各種センサを用いることで、屋内外において高い精度での位置情報測位が可能となり、位置情報を利用した様々なサービスの実用化が進んでいる。一方で、ある特定の場所におけるその場その時の状況をリアルタイムに取得することは難しい。例えば、駐車場の空きや駅、飲食店、観光スポットの混雑などの情報が挙げられる。これらの情報を所有している可能性の高いユーザにアクセスし、リアルタイムに情報の取得が可能となることで、経路選択や観光の計画を立てる際に有用となる。本研究では、情報を得たいユーザがイベントを指定した時点において、ストリームデータとして時系列に次々と発生する移動経路情報および、性別、年齢などの属性に基づいた周囲のユーザのイベントへの遭遇を予測する。そして、予測結果から情報を所有している可能性の高いユーザの絞り込みを行い、リアルタイムに効率的なユーザの検出が可能となる手法を提案する。実際の人々の行動履歴を元にした行動予測実験によって、属性を考慮した予測精度が高いことが示された。

Declarative and Continuous Query Processing for Future Event Based on People Behavior Prediction

AKINORI KAWAMURA¹ HIROKI SAITO¹

1. はじめに

GPSをはじめ、無線電波強度や各種センサを用いることで、屋内外において高い精度での位置情報測位が可能となり、このような機能を備えた携帯端末が普及している。また、それに伴い、ユーザの位置情報を用いた SNS (Social Networking Service) やナビゲーションなどの様々なサービスが実用化されている。例として、マイクロブログサービス Twitter では、その場その時の状況をツイートとしてジオタグによる位置情報とともに発信することができる。そして、ツイート内容を集約することにより、都市部におけるイベントの検出が可能である。また、位置情報を元にした SNS サービス Foursquare では位置情報履歴の共有や周辺情報の取得が可能である。また、IoT (Internet of

Things) 技術の拡大により様々な実世界情報が発信されている。このような実世界情報を蓄積した後に処理するのではなく、生成されている時系列のストリームデータを逐次処理し連続的な問い合わせ処理を行う DSMS (Data Stream Management System) に関する研究が行われている。

一方、行動予測に関する研究も活発に行われている。これらには、ユーザの位置情報履歴を元に性別や年齢などの属性の推定や、目的地予測などの研究があり、行動予測の結果に応じて広告を配信するサービスや販売促進の向上を目的とした属性に基づいたレコメンドサービスなどが普及している。しかし、これらの研究では、未来のイベントに対してユーザがその情報を収集したい際、現在の状況から起こり得る未来を予測した後、すでに蓄積された過去の位置情報やセンシング情報を検索する手法を取っているため、逐次的に生成される実世界のストリームデータを対象としたものではない。

本研究では、センサを内蔵した小型無線通信端末をユーザが所持し、ユーザが相互に実世界情報を交換し活用す

¹ 明治大学大学院先端数理科学研究科先端メディアサイエンス専攻 Frontier Media Science Program, Graduate School of Advanced Mathematical Sciences, Meiji University, Nakano, Tokyo 164-8525, Japan

る情報基盤を想定し、ユーザの行動予測に基づいた実世界のストリームデータに対する問い合わせ処理手法を提案する。提案手法は、問い合わせを行う際に、あらかじめ蓄積されたユーザの属性と行動履歴から行動予測を行い、予測結果に基づきユーザ間で交換される問い合わせメッセージの制御を行うものである。

総務省の情報通信政策研究所より2014年5月に公表された、位置情報の利用に対する意識調査 [1] によると、自分の位置情報を提供することに対し、提供する意向があるとした人は全体で65.4%であり、また自分にメリットがあるサービスのためであれば「条件付きで許容できる」を含めると70%以上が許容できると回答した。この結果から、ユーザのニーズに応じた情報を効率よく取得可能なシステムの必要性は高いと考えられる。また、京都観光総合調査 [2] の性別、年齢別の訪問地を示した表から、性別、年齢の違いによって訪れる場所に傾向があるということを読み取ることができる。このことから、同一属性のユーザは同一のイベントに遭遇する可能性が高いということが分かる。

このような背景から、本研究では、位置情報履歴から得られるイベントへの遭遇確率および性別、年齢といったデモグラフィックやユーザの意志など、ユーザ固有の属性を組みあわせることでイベントに遭遇する可能性の高いユーザの予測を行う。予測イベントに対して、リアルタイムに処理を行うことで要求される情報を所有するユーザを効率的に検出可能となる。具体的には、まずユーザの行動や属性から未来のイベントに遭遇する確率を、過去の人々の位置情報履歴の蓄積からの確率過程モデルで予測する。次に、未来のストリームデータに対する宣言的問い合わせ手法によりイベントや時間などの条件を指定した問い合わせを、未来のイベントに対する条件に一致する可能性が高いユーザ間で流通させるようにする。これらにより、ユーザが要求する未来の情報源の検出を実現する。

本論文の構成は以下のとおりである。本章に引き続き、2章では行動予測技術および実世界のストリームデータ処理などに関する先行研究を取り上げる。次に3章では、確率過程に基づく行動予測の原理および予測された未来のイベントに対する宣言的かつ連続的な問い合わせ処理の手法について説明する。4章では、行動予測による未来のイベント検出の検証実験により提案手法の実現可能性を議論する。最後に5章で本論文の内容をまとめ今後の課題を考察する。

2. 関連研究

これまで目的地推定や動線解析の研究が数多く行われている。隠れマルコフモデルによる歩行パターンからの目的地推定 [3] では、屋内環境においてある目的地に向かう人物の歩行パターンは、それぞれの歩行パターンに対応する

システムが出力すると仮定して、それらのシステムのパラメータを隠れマルコフモデルにより推定する。そして目的地が未知の歩行パターンの尤度を計算し比較することで目的地を推定している。さらに、位置情報とユーザ属性に関わる研究も行われており、位置履歴からのユーザ属性の推定 [4] では、ある空間内に設置したセンサによる、各ユーザが検出された回数を行列として表し、ユーザ属性を推定する問題を文書分類の問題に置き換えることで、SVM (Support Vector Machine) を用いた属性の推定を行っている。以上の研究では、位置情報からユーザの行動または属性を予測することを目的としている。それに対して、本研究では、属性に基づく予測結果からユーザが必要としている情報源の検出を行うことが目的である。

また、IoT (Internet of Things) 技術の拡大であらゆるものがインターネットにつながり、様々なデバイスからのセンサデータや株価、外国為替レートの変動など膨大な量のデータが発信されている。このような時系列に現れるストリームデータに対して、DSMS (Data Stream Management System) は、一度データを蓄積してから処理を行うDBMS (Data Base Management System) とは異なり、あらかじめ登録されたクエリによって連続的にストリームデータの処理を行うシステムである。さらに、DSMSを元に複数のストリームデータに対してリアルタイムに、より複雑な分析を行う技術としてCEP (Complex Event Processing, 複合イベント処理) が存在し、CEPに関する研究も盛んに行われている。連続的問い合わせに対する複数問い合わせ最適化 [5] では多数のストリーム型情報源に対する多数の連続的な問い合わせが与えられた際の効率的な実行を行う。問い合わせにおいて、同一の演算であっても実行タイミングによって全く異なる結果を生成し得る。そのため、提案手法では実行タイミングの違いによる問い合わせの参照範囲の違いを考慮し、参照範囲が近い同士の問い合わせをグループ化している。それによって、効率的な実行が可能となる複数問い合わせ最適化方式を示している。リアルタイムデータアクセス処理機構の最適化 [6] ではストリームデータを逐次的にストリーム処理する機構と、蓄積されたデータを処理する機構を兼ね備えたリアルタイム分析システムを提案している。近年、ストリームデータはリアルタイム処理に使用されるだけでなく、一度蓄積された後、分析に利用されるケースがあり、それらは別々のシステムとして存在している。この研究ではリアルタイムなソーシャルメディアデータを逐次的にストリーム処理する機構と、蓄積されたデータを処理する機構を兼ね備えたシステムを提案し、リアルタイム分析とオフライン分析との使い分けの手法を検討している。文献 [6] の手法は、すべての情報源から発せられたデータを蓄積してデータ分析に用いるのに対し、本研究では予測に基づきリアルタイムに情報源を選択する点が異なる。

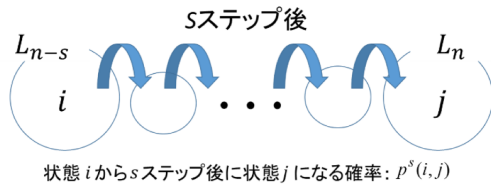


図 1 マルコフ連鎖による確率過程

Fig. 1 Stochastic process in Markov chain.

本研究では、位置情報履歴の蓄積から得られる行動予測モデルの構築と能動的情報源検出を伴う宣言的かつ連続的な問い合わせ手法を提案する。

3. イベント予測とリアルタイムマッチング

本章では、イベント予測とリアルタイムマッチングの提案手法について述べる。

3.1 手順

本研究では、1 地点を 1 つの状態とみなしたとき、過去の蓄積された人々の移動経路における状態遷移のモデル化を行う。モデルを元にイベントに遭遇するユーザを予測および検出して、情報を得たいユーザが検出されたユーザに対してリアルタイムに問い合わせ可能となる手法を提案する。提案手法では、事前にある属性を有するユーザがどのような経路を取って行動したのかという経路データを収集しデータを元にマルコフ連鎖により予測モデルを構築する。そして、情報を得たいユーザが周囲の同一属性のユーザに対して、イベントを指定した問い合わせを行う。問い合わせを行った時点において、指定したイベントに遭遇する可能性の高いユーザをリアルタイムに検出する。検出された複数のユーザを集約し、グループ化することで効率的な情報源へのアクセスが可能となる。

以降、3.2 章ではイベント予測のための確率過程を用いたモデルの構築手法について、3.3 章では未来のイベントに対する宣言的問い合わせ手法についてそれぞれ述べる。

3.2 属性に基づくマルコフ連鎖によるイベント予測手法

蓄積された過去の人々の移動経路データからマルコフ連鎖によってイベントに遭遇する可能性の高いユーザの予測を行う。マルコフ連鎖とは、次の状態が過去の状態に依存せず現在の状態のみによって決まるマルコフ性を持つ離散な確率過程である。例として、気象や株、為替など予測が難しいものに対して、マルコフ連鎖で近似することによってモデル化を行っている。提案手法では、このようなマルコフ性を適用して、人間の属性、行動といった事象からイベントに遭遇する確率のモデルを構築し、予測を行う。

ある属性 $Attr$ を有する m 人のユーザ u それぞれの移動に伴う位置 \mathbf{L} の時系列 $\mathbf{L}_u = \{L_{u,1}, L_{u,2}, \dots, L_{u,n}\} (u =$

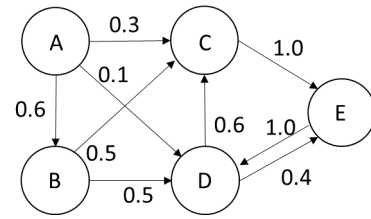


図 2 状態遷移の例

Fig. 2 Example of state transition.

$1, 2, \dots, m)$ を保持する。 m 人のユーザがある地点 L_k を通過した回数と地点 L_{k+1} を通過した回数から、連続する 2 地点間の状態遷移確率を算出する。 L_k から L_{k+1} への状態遷移確率は、式 (1) のように表せる。

$$P(L_{k+1} | L_k, Attr) = \frac{\sum^m \text{地点 } L_{k+1}(u, Attr) \text{ を通過した回数}}{\sum^m \text{地点 } L_k(u, Attr) \text{ を通過した回数}} \quad (1)$$

ある地点 L_{n-s} から任意の地点 L_n に至るまでの状態遷移確率が既知のとき、 L_{n-s} から L_n 至る確率が式 (2) により得られる。なお、変数 s は L_{n-s} から L_n までの遷移回数を表す。図 1 にマルコフ連鎖による遷移確率導出過程を示す。

$$P(L_n = j | L_{n-s} = i, Attr) = p^s(i, j) \quad (2)$$

ステップ数 s を $m+n$ と表すとチャップマン-コルモゴロフ方程式より、式 (3) が得られる。この式は、状態 i から状態 j に $m+n$ ステップで至る確率を表している。指定された状態 j およびステップ数に対して、尤度が最大となる状態 i は、状態 j に至る可能性の高いユーザが存在する地点となる。

$$p^{m+n}(i, j) = \sum_k p^m(i, k) p^n(k, j) \quad (3)$$

ある空間におけるユーザの行動履歴について、図 2 のような状態遷移を例にすると、2 地点間の遷移事前確率行列 \mathbf{M}_{pri} は以下のように表される。

$$\mathbf{M}_{\text{pri}} = \begin{pmatrix} p(A, A) & \dots & p(A, E) \\ \vdots & \ddots & \vdots \\ p(E, A) & \dots & p(E, E) \end{pmatrix} = \begin{pmatrix} 0 & 0.6 & 0.3 & 0.1 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 1.0 & 0 \end{pmatrix} \quad (4)$$

ここで、地点 E に 3 ステップで到達するユーザの予測を考える。式 (2) より、3 ステップ後に地点 E に到達する現在地点ごとの確率は、式 (5) で計算することができる。

$$P(L_n = E | L_{n-3} = i, Attr) = p^3(i, E) \quad (5)$$

よって、到達する現在地点ごとの事後確率行列 M_{post} を以下のように得ることができ、地点 E に至る最も尤度が高い (0.48) 地点 A が予測結果となる。

$$M_{\text{post}} = \begin{pmatrix} 0 & 0 & 0.18 & 0.34 & \boxed{0.48} \\ 0 & 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0.24 & 0.6 & 0.16 \\ 0 & 0 & 0 & 0.4 & 0.6 \end{pmatrix} \quad (6)$$

以上の議論に基づき、行動予測に基づく未来のイベントの問い合わせ手法の検討を行う。

3.3 未来のイベントに対する宣言的かつ連続的な問い合わせ手法

次に、未来のイベントにおける宣言的かつ連続的な問い合わせ手法について述べる。提案システムでは、CEP を応用することによって実現する。CEP の処理は以下のとおりである。全ユーザは常時通信可能であるという条件のもと、あるユーザが、条件を指定した問い合わせを行い、問い合わせを行った時点において、周囲のユーザの中から問い合わせ内に記述された条件およびイベントに合致する同一予測結果のユーザを検出する。

そこで本システムでは、ストリームデータを扱うことが可能で、予測イベントに対して宣言的かつ連続的な問い合わせが可能となるシステムを提案する。問い合わせ言語については SQL と類似の問い合わせ言語 CQL (Continuous Query Language) をさらに拡張することで実現する。3.3.1 節では提案システムの概要について 3.3.2 節では予測イベントリアルタイムマッチングについて述べる。

3.3.1 システムの概要

提案システムでは、問い合わせ元のユーザが問い合わせを行った時点においてクエリを適用し、その時刻の直前、直後におけるストリームデータの検出を行う。そのために、すべてのストリームデータの一定時間保持とリアルタイム処理を行う。また、それぞれのユーザに対するアクセスは非効率的であるため、同一予測結果のユーザをグループ化し、グループに対して、アクセスすることで効率的な情報源検出を実現する。

3.3.2 予測イベントリアルタイムマッチング

予測イベントリアルタイムマッチングでは、問い合わせを行うユーザと同一の属性をもつユーザのストリームにおいてイベントを指定した場合、3.2 で得られた結果に基づき、周囲のユーザの中からイベントに遭遇すると予測されるユーザの選択を行う。以下ではウィンドウ処理、イベント判定処理およびリンク処理を定義する。

ウィンドウ処理 CEP の特徴的な処理であるスライディングウィンドウは、あらかじめ登録された参照する時系列内のデータの個数、あるいは時間に基づく「直近

何分以内の...」といった過去の時間範囲を指定することで、元無限に流れてくるストリームデータを有限個のデータ系列として抽出することができる。本システムでは、ユーザが問い合わせを行った時点から問い合わせを開始し、指定した過去、未来の時間範囲内を参照する対象とした、能動的情報源選択を伴う宣言を行う。過去、未来に対して問い合わせを行うことで、ある時点の直前、直後の時刻でのイベントの検出を可能にする。

イベント判定処理 イベント判定処理では、ウィンドウ処理で定義した過去、未来の範囲内において生じるイベントに応じて、何らかの処理を行う際の条件を指定する。例えば、問い合わせ時に指定した時間範囲内において問い合わせ先でイベントが検出された時など、イベントに基づく条件に応じて処理を行う場合が考えられる。そこで、イベント判定処理では次々に到着するストリームデータに応じて処理を行う。

リンク処理 リンク処理とは、イベント判定処理によって判定条件を満たしたユーザのグループ化を行う処理である。リアルタイムにグループ化を行うにあたって、判定条件を満たすユーザを集約する。ネットワーク内のデータ集約による通信の効率化のためには、できるだけメッセージ数を削減する必要がある。文献 [7] では、複数の情報源に対して、最初の一つの情報源についてのみツリーの構築を行い、残りの情報源はこのつりに一おける最も近いノードに接続するというような、最適化した集約ツリーの構築手法を提案している。これを応用し、リンク処理では、ストリームにおいて先に検出したユーザを親、新たに検出されたユーザを子とする 2 分探索器などの全順序ツリーによる経路構築を行い、親が子に対して送信元を識別できる固有の情報伝達することで、最後にイベントが検出されたユーザに情報が集約される。そして、指定された時間範囲内、判定条件の下でツリーに対して深さ優先探索を行い、問い合わせ元に結果を返し、問い合わせ元は、集約された情報からそれらのユーザをグループとみなす。それによって、問い合わせ元のユーザは効率的に情報源にアクセスすることが可能となる。

4. イベント予測実験

属性に基づく行動予測結果の精度を確認するためのイベント予測実験の方法と結果について述べる。

4.1 実験方法

属性ごとのユーザの移動をマルコフ連鎖によってモデル化し、予測結果の精度を確認する実験のため、Twitter Streaming API を利用し、2017 年 3 月 9 日から 4 月 7 日までの 30 日間において、沖縄県那覇市を中心とした地域

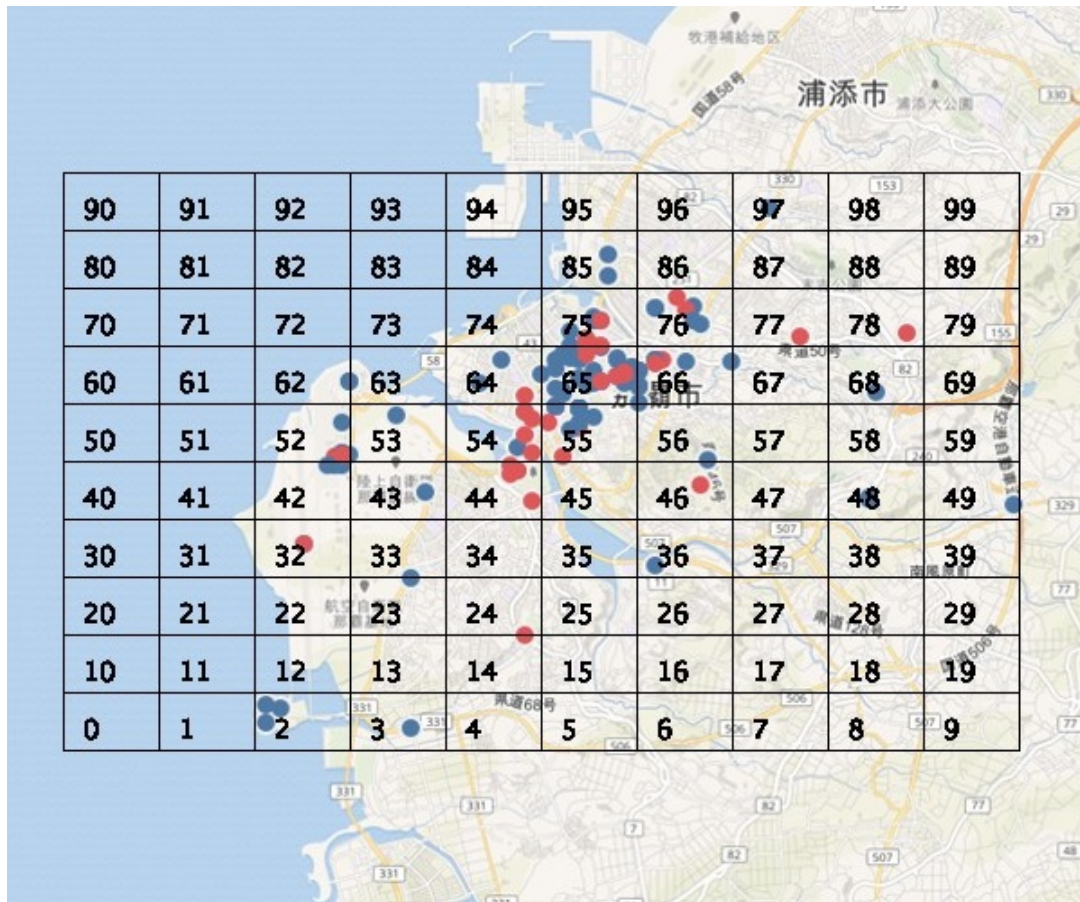


図 3 2017 年 3 月 25 日における沖縄県那覇市周辺で収集されたジオタグ付きツイート
 Fig. 3 Geo-tagged tweets in Naha, Okinawa, Japan on 25th of March, 2017.

で収集したジオタグ付きツイートをデータセットとした。ツイート収集の対象地域を 10×10 の格子状の 100 エリアに区切りエリア間の移動を状態遷移とした。また、ユーザによって n 回目にツイートした地点と $n+1$ 回目にツイートした地点が離れている可能性を考慮して、2 地点間を結ぶ直線を 100 等分した座標によって補間を行った。また、遷移は 1 日ごとに区別し、ステップ数を最大 20 とした。

収集されたツイートからユーザの属性を特定するため、文献 [8] で示された手法を用いた。これはユーザのツイート本文から属性を特徴づける複数のキーワードの出現回数を SVM で分類することにより、性別および年代の推定を行っている。これと同様に本実験でも、それぞれのユーザの過去 100 件のツイート本文から、性別については「僕」、「わたし」などの一人称をキーワードとして用いた。予備実験において性別は 96.6% の正答率であったため、推定結果を属性とした。推定の結果、ユーザ 386 人のうち、男性 252 人、女性 134 人となり、男性、女性および属性なしにおいて、それぞれ 100 人ずつをランダムに選び、モデルを作成した。図 3^{*1} は、3 月 25 日に収集されたジオタグ付きツイートが発信された場所を示しており、青が男性、赤

が女性を表す。

4.2 実験結果

それぞれの属性ごとのモデルと属性を考慮しないモデルに基づく予測結果の比較を行った実験の結果を以下に示す。表 1, 表 2, 表 3 は、属性 (性別) を考慮した予測結果と属性を考慮しない予測結果である。表中の予測地点は、首里城の存在するエリア 68 を目的地としたときの、それぞれの属性におけるステップ数ごとの尤度が最大となる地点であり、予測地点に存在するユーザが目的地に到達する可能性が高いことを示している。なお、小さいステップ数で予測された地点については、その後重複が生じないように、予測結果として出力されないものとした。表 1, 表 2, 表 3 から、男性、女性、属性なしの表に重複して現れる 17 地点のうち、エリア 77, 86, 76 を除く 14 地点は属性なしのモデルよりも、属性に基づくモデルのほうが、高い予測確率を示した。属性に基づくモデルを上回った 3 地点においても予測確率の大きな差はない。これは、ステップ数が小さく、目的地に近いことから、多くのユーザが首里城に向かっているためだと考えられる。

また、図 4 は、表 1, 表 2, 表 3 の予測地点を地図上にプロットしたものである。図 4 から、予測結果として、男

*1 Map data ©OpenStreetMap contributors.

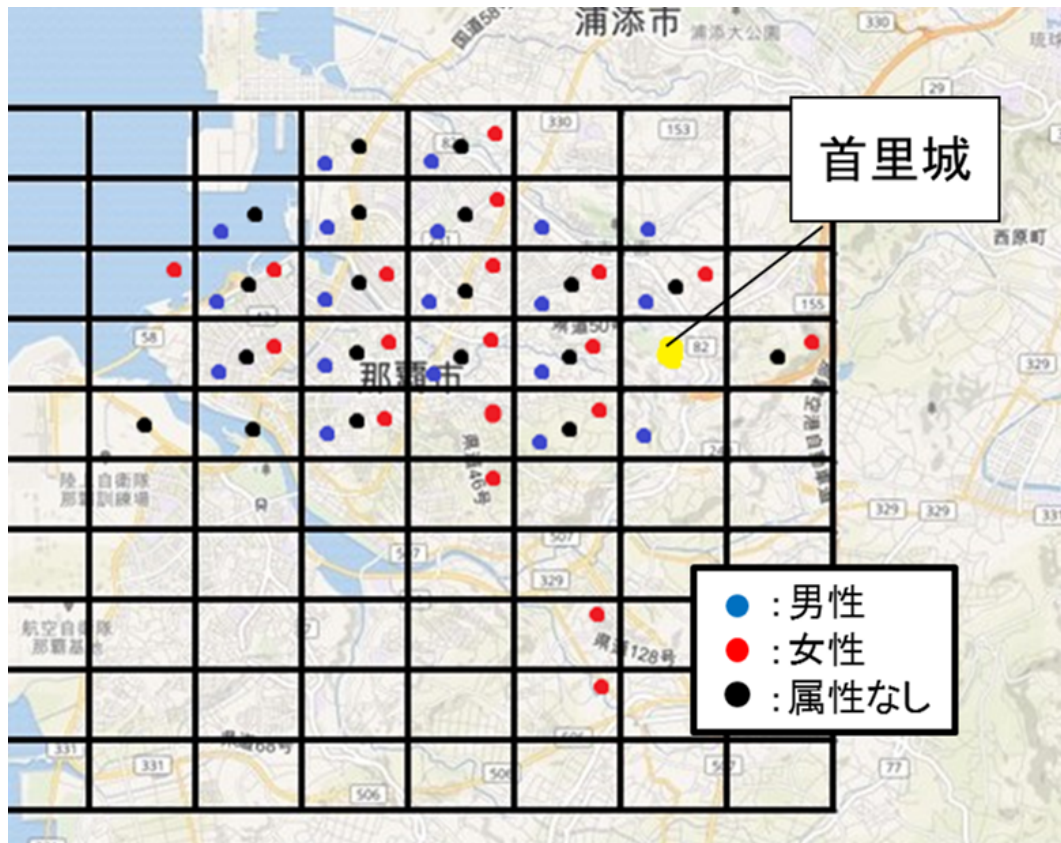


図 4 首里城付近のデータ

Fig. 4 Data around Shuri-jo castle.

表 1 属性「性別=男性」の予測結果

Table 1 Prediction result of male.

| 予測地点 | 目的地 | ステップ | 確率 |
|------|-----|------|--------|
| 77 | 68 | 2 | 0.4505 |
| 67 | 68 | 3 | 0.4295 |
| 66 | 68 | 4 | 0.217 |
| 78 | 68 | 5 | 0.3068 |
| 88 | 68 | 6 | 0.1534 |
| 76 | 68 | 7 | 0.1205 |
| 57 | 68 | 8 | 0.0972 |
| 58 | 68 | 9 | 0.0982 |
| 86 | 68 | 10 | 0.0805 |
| 65 | 68 | 11 | 0.0836 |
| 75 | 68 | 12 | 0.0799 |
| 87 | 68 | 13 | 0.0839 |
| 95 | 68 | 14 | 0.0792 |
| 85 | 68 | 15 | 0.0817 |
| 84 | 68 | 16 | 0.0788 |
| 96 | 68 | 17 | 0.0817 |
| 64 | 68 | 18 | 0.0778 |
| 74 | 68 | 19 | 0.0786 |
| 55 | 68 | 20 | 0.0755 |

表 2 属性「性別=女性」の予測結果

Table 2 Prediction result with attribute "female".

| 予測地点 | 目的地 | ステップ | 確率 |
|------|-----|------|--------|
| 77 | 68 | 2 | 0.4417 |
| 69 | 68 | 3 | 0.4479 |
| 57 | 68 | 4 | 0.2751 |
| 78 | 68 | 5 | 0.3275 |
| 46 | 68 | 6 | 0.1341 |
| 67 | 68 | 7 | 0.1797 |
| 66 | 68 | 8 | 0.1027 |
| 56 | 68 | 9 | 0.1412 |
| 86 | 68 | 10 | 0.0882 |
| 76 | 68 | 11 | 0.0955 |
| 75 | 68 | 12 | 0.0722 |
| 96 | 68 | 13 | 0.0832 |
| 17 | 68 | 14 | 0.0624 |
| 65 | 68 | 15 | 0.066 |
| 73 | 68 | 16 | 0.0574 |
| 74 | 68 | 17 | 0.0579 |
| 64 | 68 | 18 | 0.0553 |
| 27 | 68 | 19 | 0.0521 |
| 55 | 68 | 20 | 0.053 |

性ユーザのみが検出された地点が3地点、女性ユーザのみが検出された地点が5地点であり、このことから、ユーザ

属性固有の要因があると考えられる。2地点間の状態遷移確率は3章の式(1)を用いて算出しているため、ある地点

表 3 属性なしの予測結果

Table 3 Prediction result without attribute.

| 予測地点 | 目的地 | ステップ | 確率 |
|------|-----|------|--------|
| 77 | 68 | 2 | 0.4936 |
| 78 | 68 | 3 | 0.3212 |
| 86 | 68 | 4 | 0.1141 |
| 76 | 68 | 5 | 0.1303 |
| 95 | 68 | 6 | 0.0538 |
| 69 | 68 | 7 | 0.074 |
| 75 | 68 | 8 | 0.0458 |
| 96 | 68 | 9 | 0.0497 |
| 66 | 68 | 10 | 0.0387 |
| 85 | 68 | 11 | 0.0389 |
| 84 | 68 | 12 | 0.0278 |
| 67 | 68 | 13 | 0.0316 |
| 64 | 68 | 14 | 0.0259 |
| 65 | 68 | 15 | 0.0274 |
| 55 | 68 | 16 | 0.0239 |
| 74 | 68 | 17 | 0.0253 |
| 57 | 68 | 18 | 0.0158 |
| 54 | 68 | 19 | 0.0206 |
| 53 | 68 | 20 | 0.0161 |

L_k を通過した回数が 1 回であるとき、地点 L_{k+1} への遷移確率は必ず 1 となり、予測結果に影響があると考えられる。本研究の目的は、情報源の検出であるため、行動履歴が未知の領域や通過回数が極めて少数であった場合の遷移確率の計算方法は今後検討する必要がある。

5. おわりに

本研究では、ユーザの行動予測に基づいた実世界のストリームデータに対する問い合わせ処理手法を提案した。問い合わせを行う際に、あらかじめ蓄積されたユーザの属性と行動履歴から行動予測を行い、予測結果に基づきユーザ間で交換される問い合わせメッセージの制御を行う手法について説明した。位置情報履歴から得られるイベントへの遭遇確率およびユーザ固有の属性を組みあわせ、イベントに遭遇する可能性の高いユーザの予測を行う。予測イベントに対して、リアルタイムに処理を行うことで要求される情報を所有するユーザを効率的に検出可能となる。マイクロログサービス Twitter のデータセットを用いた行動予測実験により、提案手法の前提となるユーザ属性に応じた行動予測が実用的であることを示した。今後は、未来に対する宣言的問い合わせ処理方法や端末間のプロトコルの設計を行い、リアルタイム処理の手法について検討を行う予定である。

参考文献

- [1] 総務省 情報通信政策研究所, 位置情報の利用に対する意識調査 (online), 入手先 <http://www.soumu.go.jp/iicp/chousakenkyu/data/research/survey/telecom/2014/location-info.pdf> (2014) (アクセス日 2017.4.15).
- [2] 京都市, 京都観光総合調査 平成 27 年 1~12 月 (online), 入手先 <http://www.city.kyoto.lg.jp/sankan/cms-files/contents/0000202/202863/honsatsu.pdf> (2015) (アクセス日 2017.4.15)
- [3] 西村 彬宏, 森下 壮一郎, 浅間 一, 隠れマルコフモデルによる歩行パターンからの目的地推定 - 累積判別率による評価 -, 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解 106(470), pp. 43-47 (2007).
- [4] 松尾 豊, 岡崎 直観, 中村 嘉志, 西村 拓一, 橋田 浩一, 中島 秀之, 位置履歴からのユーザ属性の推定, 情報処理学会論文誌, Vol. 48, No. 6, pp. 2106-2117 (2007).
- [5] 渡辺陽介, 北川博之, 連続的問い合わせに対する複数問い合わせ最適化, 電子情報通信学会 第 14 回データ工学ワークショップ (DEWS2003), 5-C-02 (2003).
- [6] 榎 美紀, ソーシャルメディア上の情報拡散データのリアルタイム問合せ処理評価, 第 7 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2015), B1-4 (2015).
- [7] Intanagonwiwat, C., Estrin, D., Govindan, R. and Heidemann, J., Impact of Network Density on Data Aggregation in Wireless Sensor Networks, *Proceedings of The 22nd International Conference on Distributed Computing Systems (ICDCS' 02)*, pp. 457-458 (2002).
- [8] 池田 和史, 服部 元, 松本 一則, 小野 智弘, 東野 輝夫, マーケット分析のための Twitter 投稿者プロフィール推定, 情報処理学会論文誌コンシューマ・デバイス&システム (CDS), Vol. 2, No. 1, pp. 82-93 (2012).