

# 2010 年度 修 士 論 文

個人の行動履歴を利用した  
高精度目的地予測手法の構築に関する研究  
Study on High-Accuracy Destination Prediction System  
Based on Personal History of Schedule Contents

仲市 哲大  
Nakaichi, Tetsuta

東京大学大学院新領域創成科学研究科  
社会文化環境学専攻

## 目次

### 1. はじめに

- 1.1 研究の背景
- 1.2 位置情報を利用したサービスの出現と、モバイルコンテンツサービスの高まり
- 1.3 “現在位置の特定” から “目的地の予測” へ
- 1.4 目的地予測の課題
- 1.5 研究の目的

### 2. 関連研究

### 3. 目的地予測手法

- 3.1 ユーザの行動の特徴に基づく目的地予測手法の構成
- 3.2 使用するデータ
- 3.3 予測用データセット作成
  - 3.3.1 移動履歴中に含まれる滞留範囲の抽出
    - 3.3.1.1 停止判定
    - 3.3.1.2 滞留判定
    - 3.3.1.3 滞留判定閾値の更新
  - 3.3.2 スケジュール内容の定量化
    - 3.3.2.1 個人行動特徴ベクトル
    - 3.3.2.2 場所行動特徴ベクトル
    - 3.3.2.3 形態素解析
    - 3.3.2.4 TFIDF
    - 3.3.2.5 定量化済みスケジュール情報の作成
  - 3.3.3 滞留範囲とスケジュール情報の統合
- 3.4 目的地の予測

### 4. 提案手法の実証と評価

### 5. 結論と今後の課題

## 謝辞

## 参考文献

# 1. はじめに

## 1.1 研究の背景

近年、ブロードバンドを中心とする高速インターネットが急速に普及すると共に、PC や携帯電話等のデジタル機器が社会生活における基本的なツールとして広く普及している。特に最近では、第三世代(3G)携帯電話やスマートフォン、iPad を始めとするタブレット型コンピュータといった、高速な情報の通信に加え GPS センサや加速度センサ等による多様な個人行動情報の取得及び提供を可能とする高機能モバイル電子機器が世間に浸透し始めている。既に第三世代携帯電話に関しては、平成 22 年 9 月末現在で日本国内の加入契約者数が 1 億 1300 万人を突破し、88.6%の普及率を示している<sup>1)</sup>。加えて 2007 年 4 月以降、第四世代携帯電話からの緊急通報について発信者の位置情報等を通知する機能を備えることが電気通信事業者に義務付けられたことから<sup>2)</sup>、GPS センサ付きの携帯電話が標準規格となって市場に流通するようになった。これら情報通信インフラの整備と高機能モバイルの普及に伴い、電子機器を所有するユーザ(以下ユーザ)が自身のあらゆる行動を空間情報、特に緯度経等の位置情報と共にリアルタイムで補足し、その内容を電子化することが可能になってきている。

また、取得した情報を取り扱う手段も多く現れるようになった。特に 2005 年から日本でも利用が可能となった Google Map や Google Earth が、多くの人々を位置情報の世界に慣れ親しませてきた。またこれらを扱うために必要な設計図である KML(Keyhole Markup Language)の公開によって、人々は簡単に自身の行動や位置情報を加工し、世界中に発信できるようになった。これを大きな契機として、それまで専門的な領域とされてきた位置情報がウェブサービスを通じて人々により一層身近な存在となり、様々な企業が立ち上がった。ユーザはこれらの企業に自身の情報を提供し、その代償として自らが求める他の情報や便利なサービスを受けてきた。

政策面でもこの流れを踏まえ、2007 年 5 月には国会で「地理空間情報活用推進基本法(2007 年 5 月 30 日法律第 63 号)」が成立した。これは国や自治体が所有する基盤地図情報を国民に公開することで誰もが位置や場所をキーとして必要な情報を容易に検索・統合・発信・利用し、また位置をキーにして多様な情報を統合することで新たな価値を生み出すことができる行動地理空間情報活用社会を目指している。また 2008 年 4 月にはこの社会の実現に向けた行動を詳しく取り決めた「地理空間情報活用推進基本計画」が閣議決定された。この計画では、国や自治体が所有する基盤地図情報を国民に公開することで、誰もがいつでもどこでも必要な地理空間情報を利用可能にし、また高度な分析に基づく的確な情報を入手し行動できる行動地理空間情報活用社会を目指している。

この様に、「モバイル機器による位置情報取得手段の普及」「位置情報加工及び提供・公開手段の拡大」「各関係法整備の充実」というあらゆる面からの追い風により、位置情報を用いたサービスは今後より一層高まりを見せていくことが予想される。

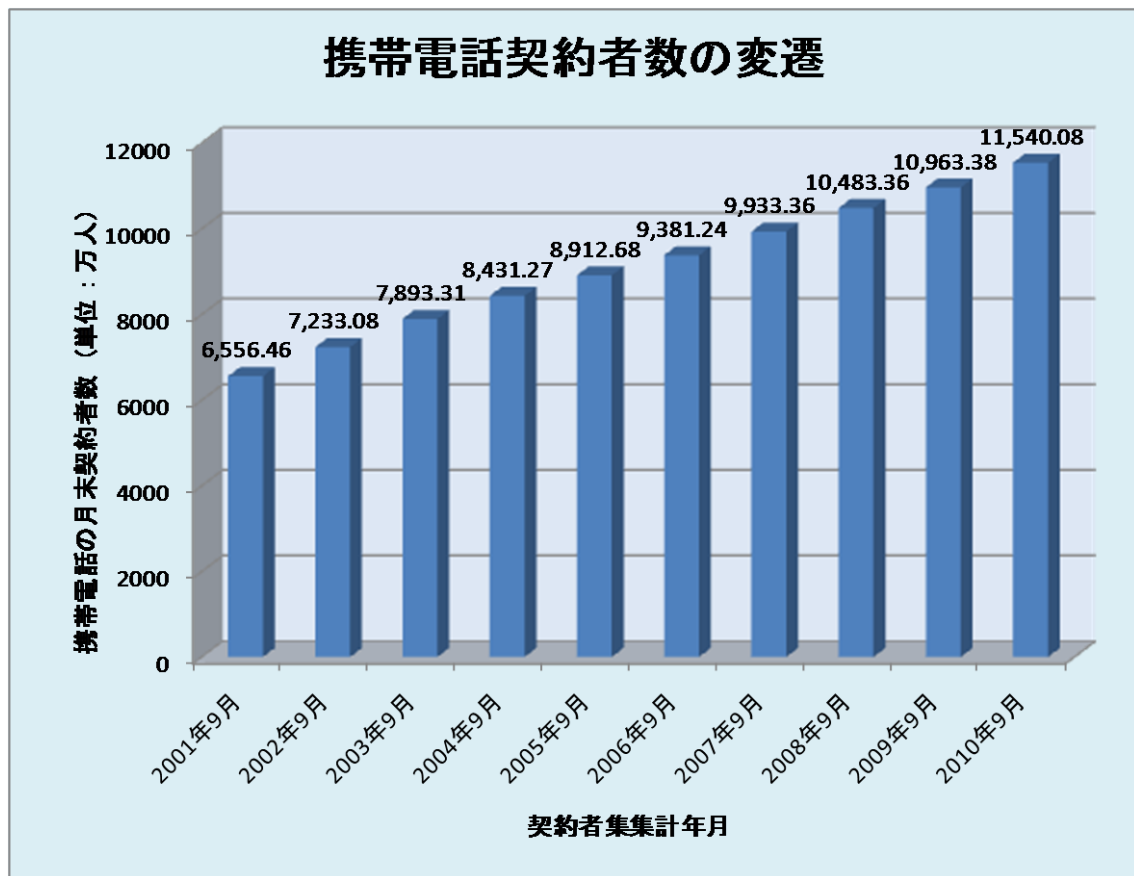


図 1 携帯電話契約者数の変遷

## 1.2 位置情報を利用したサービスの出現と、モバイルコンテンツサービスの高まり

モバイル機器によって取得される位置情報を利用したサービスの一例として、NTT ドコモが同社の携帯端末向けに提供しているサービスに i コンシェル<sup>3)</sup>がある。現在の日時に加え、GPS による測位や交信中の基地局から端末の現在位置を割り出し、付近の鉄道運行情報や道路交通情報・気象情報など、状況に合致した情報をユーザに提供する。i コンシェルのユーザ数は 2010 年 3 月末で約 420 万人<sup>4)</sup>であり、これらユーザに情報を提供する一般企業は 500 社を超えるなど、大きな市場が形成されつつある。また NAVITIME JAPAN<sup>5)</sup>では、ユーザが向かう目的地までの移動ルートを案内するナビゲーションサービスを提供している。GPS で取得した位置情報に基づく現在位置をモバイル画面の地図上に表示し、音声ガイドダンスと共に目的地までのルートを案内する。

上記の様な情報提供サービスとは別に、コミュニケーションサービスにも個人の行動情報を付与する仕組みが増えつつある。例えば twitter<sup>6)</sup>は、個々のユーザが「ツイート」と呼ばれる 140 字以内の文をウェブ上に投稿し、投稿された文は他のユーザが自由に閲覧できるといった一種のコミュニケーションツールである。このツイートをから投稿する際、ユ

ーザ自身の現在位置を添付して投稿することができる。また foursquare<sup>7)</sup>は位置情報を利用したソーシャルネットワーキングサービス(SNS)の一種である。現在位置を不特定多数のユーザに公開し、公開した回数によって得点やバッジと呼ばれる勲章を受け取る。

位置情報を利用したゲームも登場している。コロニーな生活☆Plus<sup>8)</sup>は与えられた仮想都市(コロニー)を仮想通貨(プラ)の使用によって成長させるゲームである。仮想通貨の取得に位置情報を利用する。滞在地毎に次の滞在地の位置座標との距離に応じて仮想通貨を得るという仕組みである。ケータイ国盗り合戦<sup>9)</sup>は、ユーザが滞在した先の現在位置を取得することで現在位置が所属する地域を「統一」し、これを繰り返しながら日本全国の全地域を統一するゲームであり、加入者数は2010年12月24日時点で70万人を超える<sup>10)</sup>。

加えて、これらサービスが属するモバイルコンテンツ市場そのものも年々成長の一途をたどっており、2009年のモバイルコンテンツ市場規模は5525億円にも上る<sup>11)</sup>。その中で地図情報やナビゲーション等の位置情報に基づく情報を提供する交通情報市場に関しては241億円、モバイルゲーム市場に関しては884億円とモバイルコンテンツ市場全体から見れば占める割合はそれぞれ4%、16%と未だ少ない。しかしながら両者とも毎年堅調に成長を遂げており、位置情報サービスとモバイルコンテンツ全体の勢いは今後益々加速していくと考えられる。

表1 位置情報を用いた主なサービス例 <sup>4)6)7)10)12)13)</sup>

サービス名	事業者	サービス開始時期	加入者数
iコンシェル	NTTドコモ	2008年11月	420万人(2010年3月) <sup>4)</sup>
NAVITIME	NAVITIME JAPAN	2000年3月	400万人(2009年9月) <sup>12)</sup>
Twitter	Twitter, Inc.	2006年7月	1億7500万人 (全世界 2010年9月) <sup>6)</sup>
Foursquare	Dennis Crowley, Naveen Selvadurai	2009年3月	500万人(2010年12月) <sup>7)</sup>
コロニーな生活 ☆Plus	株式会社コロプラ	2005年5月	100万人(2010年4月) <sup>13)</sup>
ケータイ国盗り合戦	株式会社マピオン	2008年4月	70万人(2010年12月) <sup>10)</sup>

表2 モバイルコンテンツ市場の市場規模

モバイルコンテンツフォーラム

「モバイルコンテンツ関連市場に関する市場規模調査」より引用

(単位：億円)

	2006年	2007年	2008年	対前年比	2009年	対前年比
着メロ系市場	843	559	473	85%	402	85%
着うた系市場	759	1,074	1,190	111%	1,201	101%
(内訳) 着うた市場	(522)	(568)	(483)	85%	(432)	89%
(内訳) 着うたフル市場	(237)	(506)	(707)	140%	(769)	109%
モバイルゲーム市場	748	848	869	102%	884	102%
装飾メール市場	55	116	171	147%	228	133%
電子書籍市場	69	221	395	179%	500	127%
リングバックトーン市場	29	87	110	126%	115	105%
占い市場	158	182	200	110%	191	96%
待受系市場	248	227	229	101%	226	99%
きせかえ市場	3	23	64	278%	99	155%
天気／ニュース市場*1	63	73	78	107%	97	124%
交通情報市場*2	125	164	206	126%	241	117%
生活情報市場*3	45	54	77	143%	121	157%
アバター／アイテム販売(SNS等)*4	5	60	157	262%	447	285%
動画専門市場*5	24	36	62	172%	112	181%
芸能・エンタテインメント系市場*6	187	195	201	103%	241	120%
メディア・情報系市場*7	74	77	66	86%	66	100%
その他モバイルコンテンツ市場	231	276	287	104%	354	123%
モバイルコンテンツ市場合計	3,666*	4,272*	4,835	113%	5,525	114%

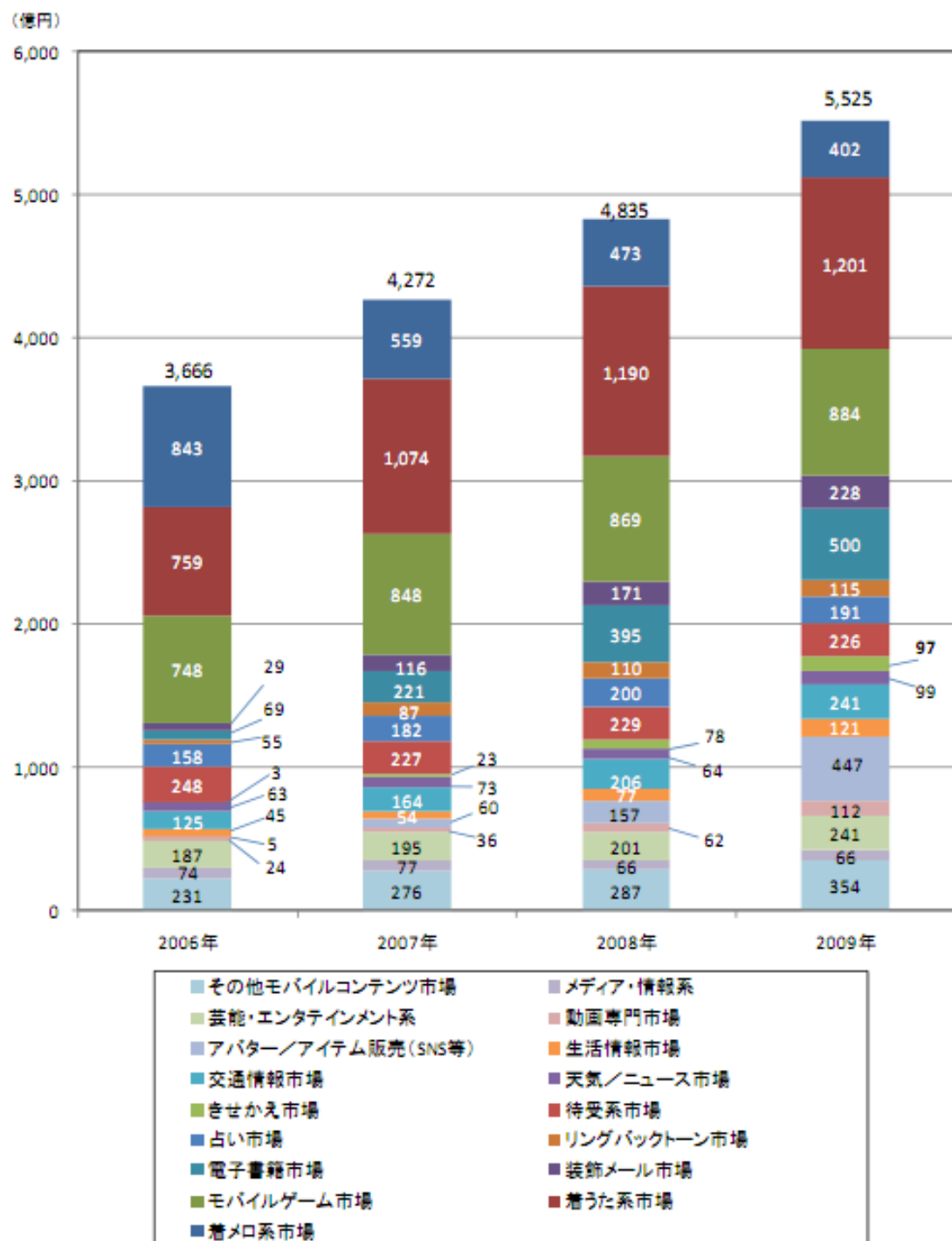


図2 市場規模内訳

モバイルコンテンツフォーラム

「モバイルコンテンツ関連市場に関する市場規模調査」より引用



図3 i コンシエルのサービス概要

NTT ドコモ HP より引用

<http://www.nttdocomo.co.jp/service/customize/iconcier/function/>



図4 ナビゲーションサービスの概要

NAVITIME ホームページより引用

<http://corporate.navitime.co.jp/index.html>



### 1.3 “現在位置の特定” から “目的地の予測” へ

このように、位置情報は最早サービス提供に欠かせない重要な要素の一つとなりつつある。現状ではユーザの“現在位置”に対して何らかのサービスが付随するものが一般的であるが、もしこれから向かおうとしている“目的地”が事前に予測され、予測された目的地の属性やその行動内容があらかじめ判明すれば、これまで以上に便利かつ実用的なサービスが提供できると考えられる。

具体的には、個人の行動履歴から行動の特徴を学習(パーソナライズ)し、移動経路や目的地、さらには行動目的を先読みすることで、ユーザが興味を持っている施設やスポットを案内したり、また事前に目的地までの移動手段を複数検索して確実に予定時間内に目的地へとたどり着けるルートを推薦してくれたりといった、ユーザの生活行動を支援するサービスを提供することが可能になる。

またユーザだけではなく、サービスを提供する企業側にとっても大きなビジネスチャンスが発生する。その最たる例の一つがローカルアドミニストレーション(Local Administration, 以下ローカルアド)である。ローカルアドはこれまでの不特定多数に対して発信する宣伝広告とは異なり、ウェブ上の検索結果やユーザの滞在場所、生活エリア及び興味を持った場所へ向けてピンポイントで発信する広告である。関心を持ちそうなユーザに対してのみ広告を打つため、効率的かつ効果的な広告宣伝活動が可能となり、すでにアメリカでは Chitika<sup>14)</sup>社が、日本でもシリウステクノロジー社<sup>15)</sup>がサービス提供を開始している。目的地の予測とローカルアドとが連動すれば、ユーザは事前に目的地周辺に存在する店舗情報を得ることが可能になる。そのため当該店舗に対する興味関心が早くから起こり、これまで以上に広告効果が高まるものと予想される。

これらは目的地予測によって可能となるサービスのほんの一例であり、将来的にはさらに多岐にわたるサービスが考案・提供されると考えられる。



図5 ローカルアドミニストレーションサービス

Chitika.Inc より引用 <http://chitika.com/>



図 6 目的地予測により可能になるサービスの例

## 1.4 目的地予測の課題

1.3 や図 6 において示したような新たなサービスを実現するには、ユーザが向かう目的地の予測、それも高い確率で目的地を予測できる手法が不可欠である。目的地予測はこれまで多くの研究者によって手法が開発研究されてきた。その多くが、ユーザの過去の行動履歴、特に移動履歴からユーザが訪れたことのある場所を特定し、それら場所同士の関係性に基づいて、これからユーザが向かうであろう目的地を推測するというものである。移動履歴を取得する際には、一般に GPS センサが用いられる。

しかしながら、既存の研究における目的地予測結果と実際に訪れた場所の一致率は十分ではなく、直近の研究においてさえ、その精度は 68.2%<sup>12)</sup>と高精度とは言えないものである。

この理由として、GPS による測位では現在位置(緯度・経度・標高)と位置取得時の時刻しか判明せず、具体的な行動の内容(ミーティング、旅行など)や場所を示す意味情報(セマンティクス)は、位置情報を基に他の情報から別途付与する必要がある。そのため単なる移動の履歴として座標列を得ることはできても、具体的に立ち寄った場所や目的地(到着地)、そこでの行動までを推定することはできない。人々がある場所に滞在するのは基本的に何らかの行動をそこで行うためであり、その行動の内容によって訪れる場所が決定される事が多い。そのため行動履歴から滞在した場所の位置情報だけを抽出し、滞在所間関係を見ただけではユーザの移動を予測するには十分ではなく、これに加え「ユーザは次にどのような行動を取るか」「この行動を行う場合にはどの場所へ向かうか」といった行動の特徴を把握した上で目的地を予測する必要がある。

## 1.5 研究の目的

本研究では、新たな位置情報サービスの創出に必要不可欠な、より高精度な目的地予測手法の構築を目指す。その際に上記の問題点を克服するため、移動履歴からユーザの滞在所を抽出する従来手法に加え、ユーザの行動特徴を表す情報としてユーザが日々つけている電子スケジュール情報を用いることで、行動の特徴を正確かつ自動的に把握し、目的地予測に利用することとした。スケジュール情報には、行動の内容(タイトル)、目的地、開始時と終了時の時間など、個人の行動を示す情報が多く含まれている。これらの情報を基にユーザの行動の特徴、及びユーザが目的地に対して有している特徴を計算し、これに GPS データからなる移動履歴を組み合わせることで、ユーザの現在位置や時刻、そして予測される行動に最も適した目的地の推定を行う。

## 2. 関連研究

本章では、従来の目的地予測システムに類する代表的な研究をいくつか取り上げ、その内容を解説・整理する。その上で位置情報と行動特徴を結びつけて行う目的地予測の必要性を述べ、既存研究のシステムと本研究で提示する目的地予測システムの差異を明確化する。

従来の手法では、その多くが GPS 位置情報履歴からユーザが過去に目的地として滞在した場所を特定し、その目的地間の遷移関係に基づいて目的地を予測する手法を提案している。滞在場所は主にクラスタリングと呼ばれる分析によって特定される。クラスタリングとは、分類の対象となるデータの集合を、ある基準に基づいてクラスと呼ばれる集団に分割する手法である。目的地予測においてクラスタリングを用いる際は、データの集合が GPS によって取得した座標、分割されたクラスが滞在場所即ち過去に訪れた目的地となる。

Ashbrook<sup>16)</sup>はユーザが長時間居たところ、つまり GPS で取得した座標が数多くかつ一定範囲内に集中してプロットされている場所が各移動の目的地であると考え、この場所を k-means 法によって特定している。西野<sup>17)</sup>らも同様の考えに基づいて、DBSCAN と呼ばれるクラスタリング手法により目的地を特定しようと試みている。k-means 法は一般的なクラスタリング手法の一つである。計算が簡単でシステムを構築しやすいが、局所最適解を求める手法のためランダムに決定された初期条件によって最終結果は大きく影響される。DBSCAN 法は、データ集団中のデータ密度が高い集団をそれぞれひとつのクラスとして取り出し、いずれのクラスにも属さない点をノイズとして除去するクラスタリング手法である。この手法では k-means 法のようにただ座標同士の重心を計算するだけでなく、点の集まり具合(密度)にも着目してクラスを生成するため、座標が密集してプロットされている箇所、つまり長時間留まっていた場所をより正確に検出することが可能となる。またクラスタリングによって求めた目的地ではそれぞれ「A にきたら次は必ず B に向かう」「C に到着したら次はおおかた D へ向かっている」等の相互間の遷移関係が計算される。この遷移関係の計算に関しても様々な手法が提案されており、Ashbrook はマルコフモデル、西野らは時系列マイニングにより計算を行っている。マルコフモデルは、時間経過に伴って変化する確率変数を表すモデルである。ここでは、次に訪れる場所は現在滞在している場所のみで決定され、その前の場所やさらに過去の場所は全く無関係であるとした上で目的地間の遷移関係を計算している。時系列マイニングによる遷移関係の計算では、場所間の遷移に時刻情報の影響も加味している。以上の手法によって抽出された目的地と計算された遷移関係に基づいて、現在の滞在場所から次の目的地を予測することが可能となり、上記の研究以外にも多くの手法を適用した研究がこれまでなされている。

しかしながら、山田<sup>18)</sup>はこれら単に目的地間の遷移関係に基づいた予測手法に対し複数の問題点があると指摘する。そのうちの一つは、高遷移確率目的地のみの予測である。自

宅や通勤先・通学先といった移動の開始地点となる滞在場所から目的地予測を行う場合には、常に遷移確率が最大の目的地が予測結果として提示されてしまい、その他多くの予測候補地について誤った予測を行ってしまう。例えば、自宅からの遷移確率最大の目的地が大学であるとする、自宅に居る場合の次の訪問地は必ず大学になってしまい、通学途中で本屋に立ち寄り、食事のためにレストランを訪れるといった可能性を全て無視する結果になる。また他の問題点として、ユーザが初めて訪れた場所からでは次の目的地が予測できないという点が挙げられる。これは初訪問の場所には他の場所との遷移関係が存在しないためであり、例えばユーザが帰宅途中に新しくできた映画館で映画を見てから自宅に帰ったとする。この場合、この映画館から他の目的地への遷移関係が存在しないため、映画館の次の目的地として自宅を予測することは不可能になる。

この問題点を解消するために、山田ら<sup>18)</sup>や Zhon ら<sup>19)</sup>は移動経路履歴に基づいて目的地を予測する手法を提案している。GPS 位置情報履歴から直接目的地を求めるのではなく、移動した道筋即ち移動経路をクラスタリングによって抽出し、予測に用いている。これにより、ユーザの現在位置からその地点を過去に通過した移動経路を特定し、当該移動経路上で一定時間滞在した場所、及びそこへ至るまでの経由地を予測結果として提示している。また初めて訪れた場所における次回目的地の予測に関しても、既知の経路上にユーザの位置が復帰した時点でなら可能になった。

だが、これらの手法でも先に述べた Ashbrook の例と同じくユーザの移動履歴の中から位置情報のみに基づいて過去の滞在場所を類推した上でその滞在履歴から目的地予測を予測しており、予測結果の的中率も概して高いとは言えない。

移動履歴だけでなく、他の情報を用いて滞在場所の詳細な情報を取得し、目的地予測に利用した研究も存在する。Liu<sup>20)</sup>らはクラスタリングによって滞在場所を特定した後、滞在場所が属する番地名を地図データから取得する。次にその番地に存在する建物名を電話帳から検索し、ヒットした建物をユーザが訪れた詳細な滞在場所と見なす。さらにその建物の業種カテゴリから、ユーザがその時行っていた行動の内容を類推する。この様にして位置情報とテキスト情報という性質の異なる情報同士を組み合わせ、目的地予測に利用した。

本研究では、移動履歴（位置情報履歴）に加え、ユーザの行動の内容を表す行動特徴情報、それもユーザが日々記録している電子スケジュール情報から抽出した行動特徴を目的地予測に用いる点で、従来手法と大きく異なる。先にも述べたとおり、スケジュール情報には、行動の内容、目的地、開始時と終了時の時間といった個人の行動を示す情報が豊富に含まれている。従ってスケジュール情報を解析することで、ユーザがどのような行動を取っているかが如実に明らかになる。行動の内容はユーザが目的地を選択する際に大きく関与すると考えられるため、これらを従来の位置情報に基づく目的地予測に組み込むことで、予測精度の向上が見込まれる。位置情報に他の情報を組み合わせる目的地予測手法は、Liu らを始めいくつかなされているが、スケジュール情報由来の行動内容を付与する手法はこれまで皆無であり、この点においては本研究が新たな視点を開いたといえる。

### 3. 目的地予測手法

#### 3.1 ユーザの行動の特徴に基づく目的地予測手法の構成

本研究で提案する目的地予測手法(以下, 提案手法と呼称)では, 図9に示す手順に沿って, まずユーザの行動履歴から予測用データセットを作成し, このデータセットに基づいてユーザが今居る現在地点から次に向かう目的地の予測を行う.

まず予測の前段階として, 移動履歴(図8のI-1. 以後の記号も同様に図8に対応)とスケジュール履歴(I-2)を組み合わせ, 予測用のデータセットを作成する. このデータセットでは, ユーザの移動履歴から取得した位置情報とそこに滞在していた時間の情報, またスケジュール履歴から取得した場所名・行動内容(タイトル)及び行動が実施された時間の二種類の履歴情報を使用する.

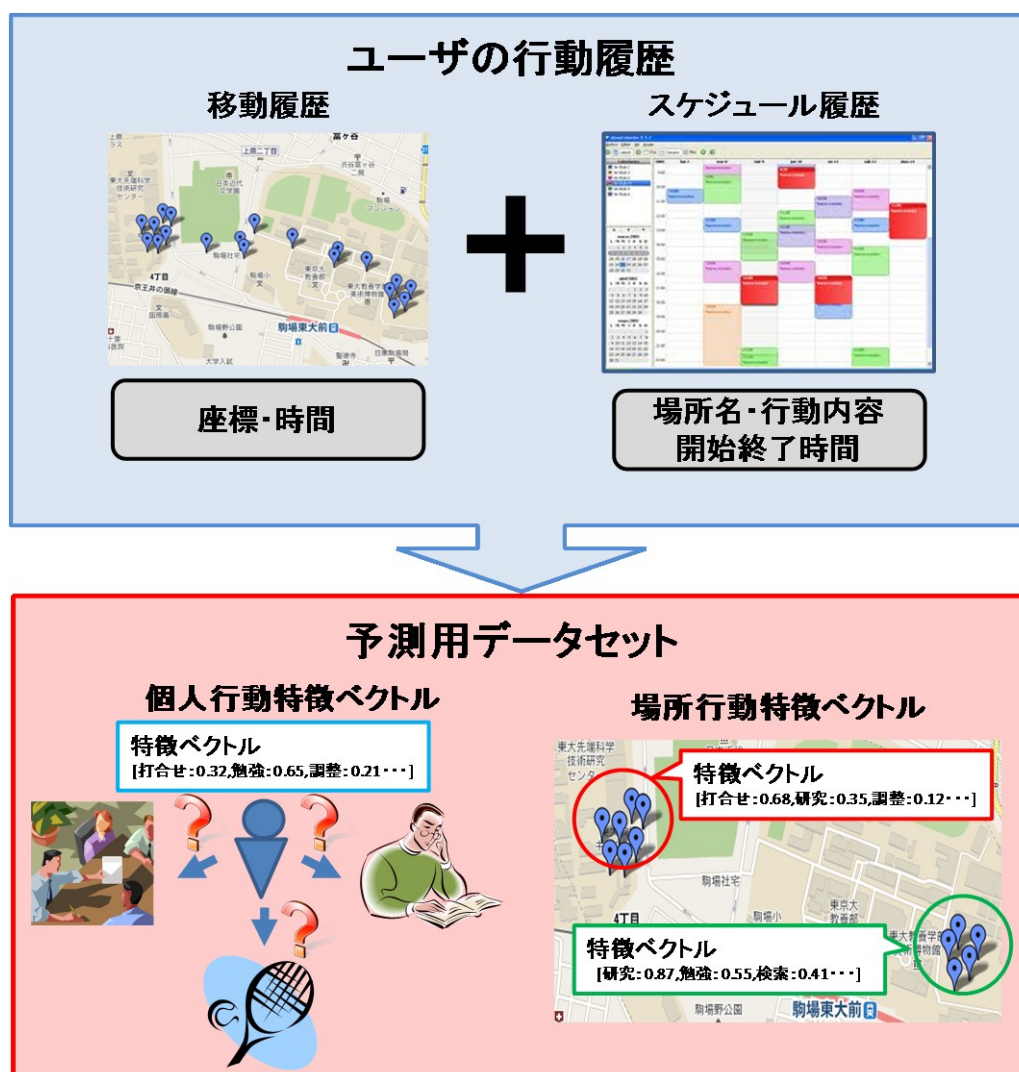


図7 行動履歴から作成される予測用データセットの概念

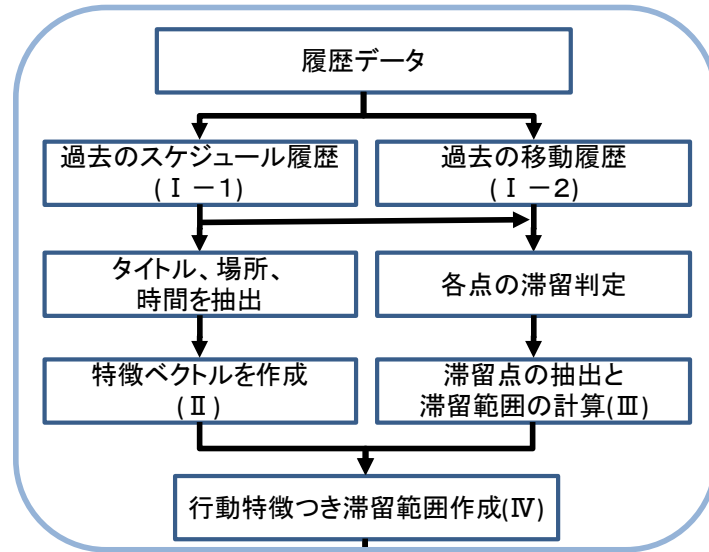
本研究の提案手法において取得を目指す目的地とは、「ユーザが何らかの目的を持って訪問し、目的を完遂するために一定時間滞留した範囲」とし、本研究では以後これを『滞留範囲』と呼称する(Ⅱ)。また当該目的地で実施される行動の内容を表す情報として「ユーザがある滞留範囲において行った行動の内容」を計算し、以後これらを『場所行動特徴ベクトル』と呼称する(Ⅲ)。位置情報とテキスト情報という性質の異なる情報は、両者が有している時間情報を用いて統合する(Ⅳ)。これにより数値的な位置情報に対して、その場所の意味を示す情報が付与可能となる。この位置情報とテキスト情報の統合結果を本研究では『行動特徴つき滞留範囲』とする。またこれら場所に付随する行動の特徴とは別に、ユーザ自身の行動の特徴も必要となる。すなわちユーザの全てのスケジュール履歴から行動の内容を取得し、「ユーザがこれまで行った全ての行動の特徴」加えて「ユーザがこれまで行った時間帯毎の行動の内容」を計算して、これらをそれぞれ『個人行動特徴ベクトル』及び『時間帯別個人行動特徴ベクトル』と呼称する。この個人行動特徴ベクトルと行動特徴つき滞留範囲の二つをもって、予測用データセットとする。

このデータセットを元に目的地の予測を行う。まずユーザの現在時刻及び現在位置を取得する(A)。次いでユーザの個人行動特徴ベクトル及び時間帯別個人行動特徴ベクトルから(B)、現時点でユーザが取るであろう行動を決定し、また現在位置より一定範囲内における行動特徴つき滞留範囲を全て抽出する(C)。最後に抽出した滞留範囲の中から、ユーザがこれから行くと予想された行動が最も実行される割合の高い場所を、現時点で最も確からしい目的地として提示する(D)。

以上が提案手法の一連の流れである。



## 予測用データセット部分



## 目的地予測部分

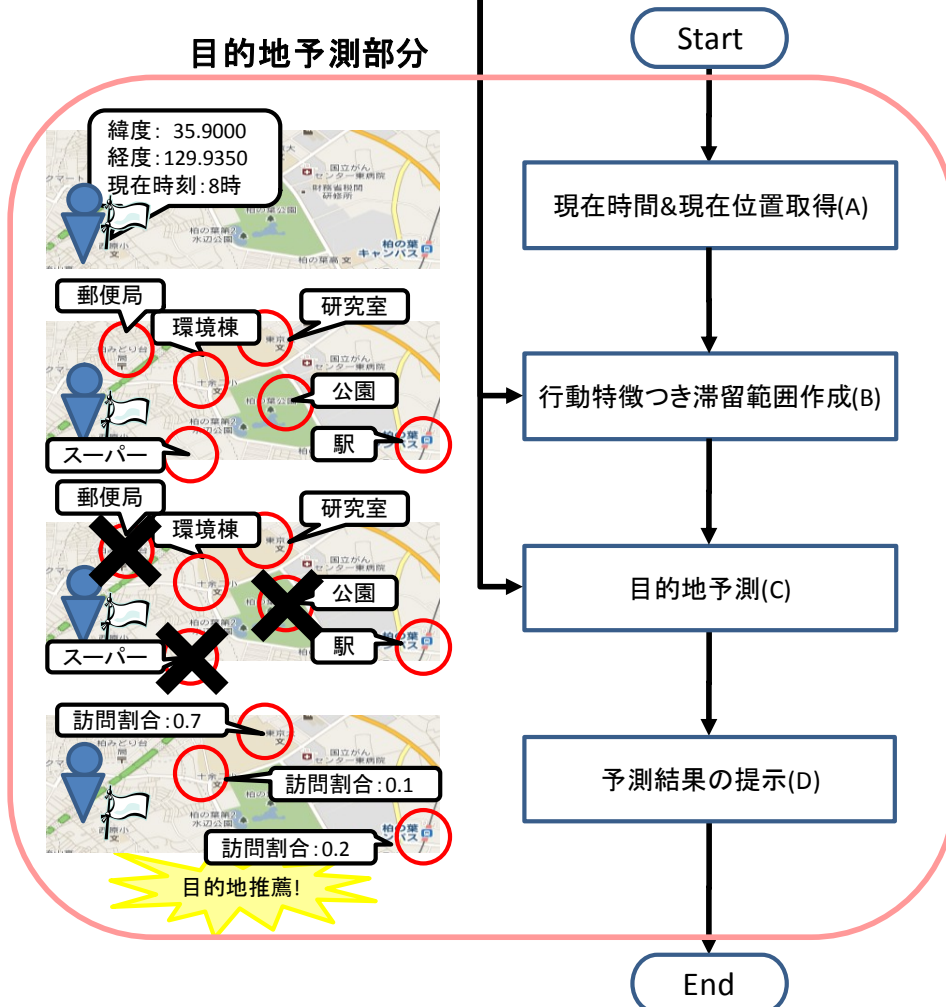


図 8 提案手法の全体概要



### 3.2 使用するデータ

提案手法では、目的地予測の際の入力データとして、ユーザの「移動履歴」と「スケジュール」を利用する。移動履歴は GPS によって取得された緯度経度からなる座標情報を用いる。本研究では、移動履歴の取得に、GPS ロガー及び GPS センサ付きのスマートフォンを使用した。GPS ロガーには QSTARTZ 社製の「BT-Q1300S GPS Sports Recorder」を、GPS センサ付きスマートフォンにはソニーエリクソン社製の NTT ドコモの第三世代移動通信システム端末である「SO-01B(Xperia)」を使用した。それぞれの機器の詳細な仕様は図 12 に示す。またスケジュールに関しては、データ取得の容易さ及び加工・解析の容易さから、パソコンや携帯電話で用いられるデジタルスケジュールから情報を取り入れることとした。特に本研究では iCalendar 形式のスケジュールフォーマットを用いた。iCalendar 形式は一般に普及しているスケジュール管理ソフトで共通に利用可能な形式であり、Microsoft Outlook や Google Calendar, Mozilla Thunderbird 等でも用いられている。また様々なプログラミング用ライブラリも公開されているため、システムの構築が非常に容易である。本研究では“iCal4j<sup>21)</sup>”と呼ばれる JavaAPI を用いて、スケジュールデータの加工・解析を行った。

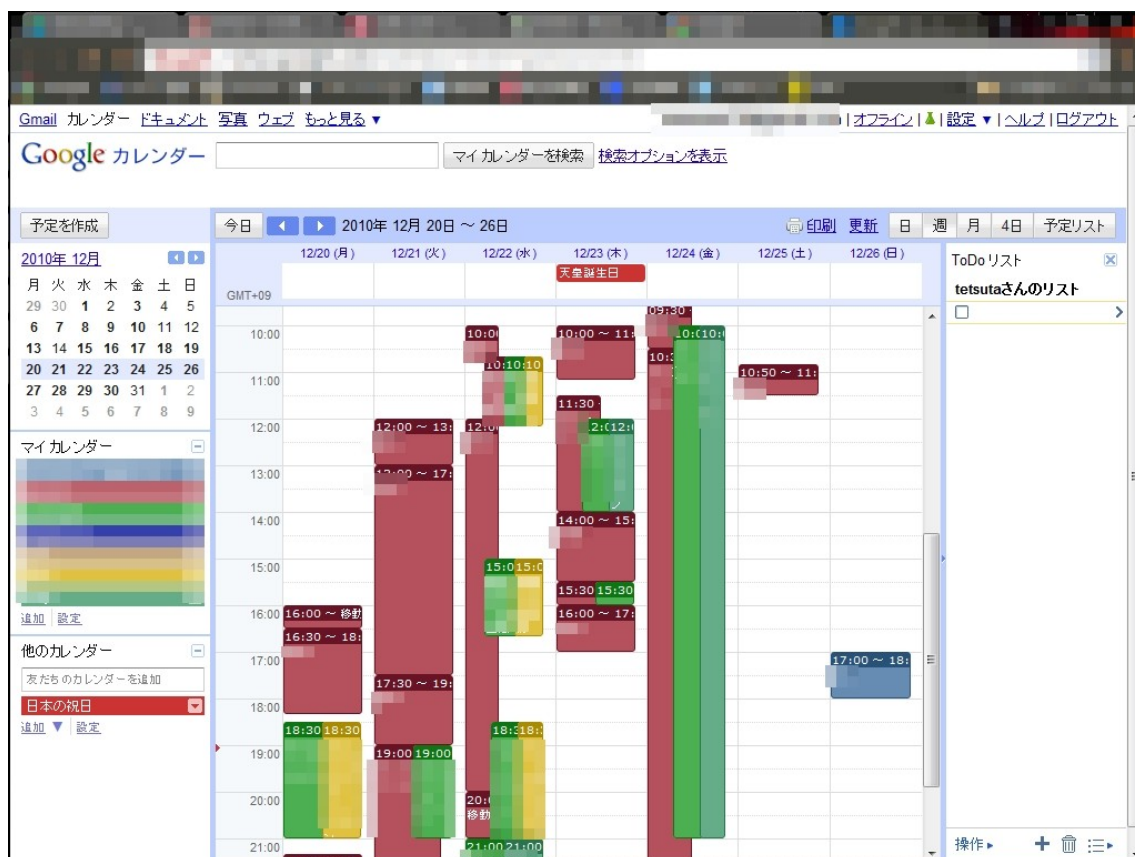


図 9 iCalendar を採用したスケジュールアプリケーション（図は Google Calendar）

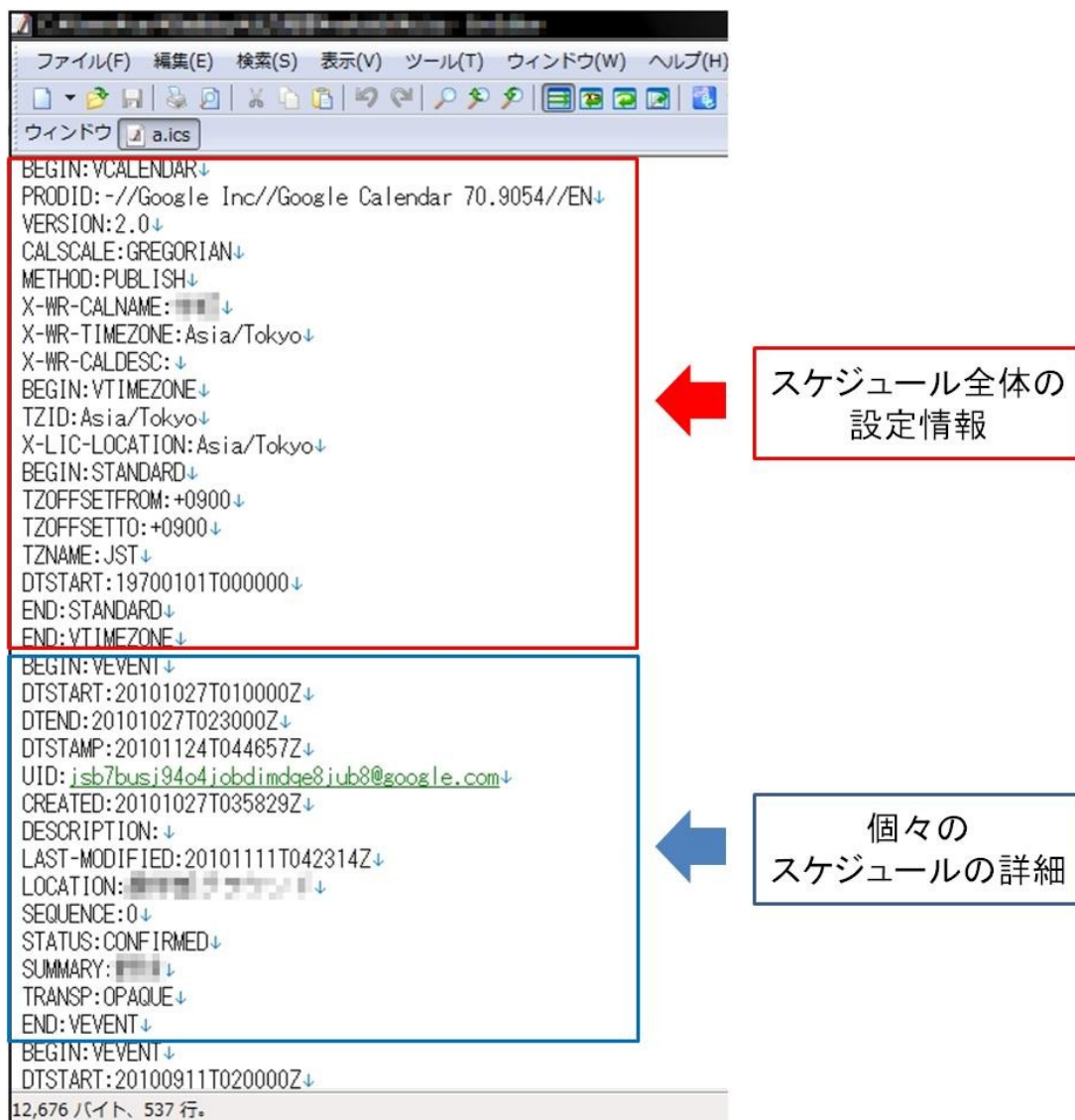


図 10 iCalendar のデータサンプル



図 11 GPS ロガー及び GPS センサ付きスマートフォンの外観と仕様

### 3.3 予測用データセット作成

本章では目的地予測に用いる、予測用データセットの作成について述べる。

#### 3.3.1 移動履歴に含まれる滞留範囲の抽出

移動履歴には、ある場所に留まっている際に取得された座標もあれば、移動しているときに取得された座標も含まれる。目的地予測においては、ユーザが目的を持って滞在していた場所を選ぶ必要があるため、様々な座標の集合である移動履歴からある場所に留まっている時の点のみを正確に抽出しなければならない。本項では移動履歴を移動中の点と停止中の点に分類し、その上で停止中の点を単なる一時的な停止かそれとも目的をもった停止即ち滞留であるかを判定し、さらに滞留と見なされた点を一定範囲内で集約し、面の要素を持つ滞留範囲として整形する手順について説明する。

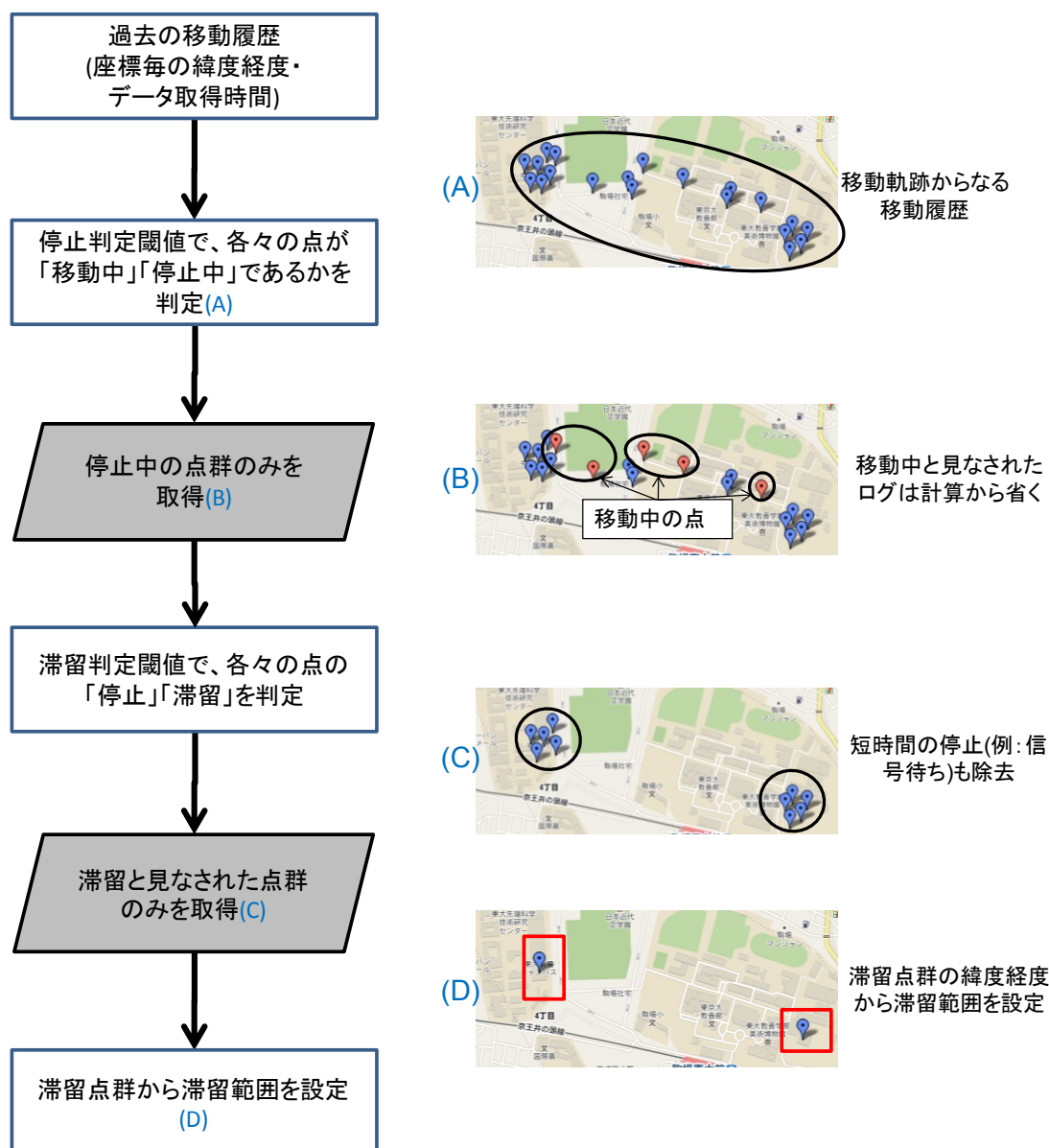


図 12 移動履歴に含まれる滞留範囲抽出の流れ

### 3.3.1.1 停止判定

始めに移動履歴を停止点と移動中の点に分類する。

移動履歴には座標  $P_1, P_2, \dots, P_m$  まで  $m$  個の点が含まれている。この中からまず、最初の点  $P_1$  から  $P_2, P_3, \dots, P_n$  と  $n$  番目までの点に着目し、これら  $n$  個の点群において  $x, y$  座標の重心を求める。次に求めた重心と各点との距離を算出し、これらの標準偏差  $d$  を導く。この  $d$  が停止判定用閾値  $T$  よりも大きければ  $P_n$  は移動していると見なし、それ以外ならば停止中の点と見なす。 $T$  の初期値は 0 であり、これは 1 日の最後に更新される。閾値更新については後述する。

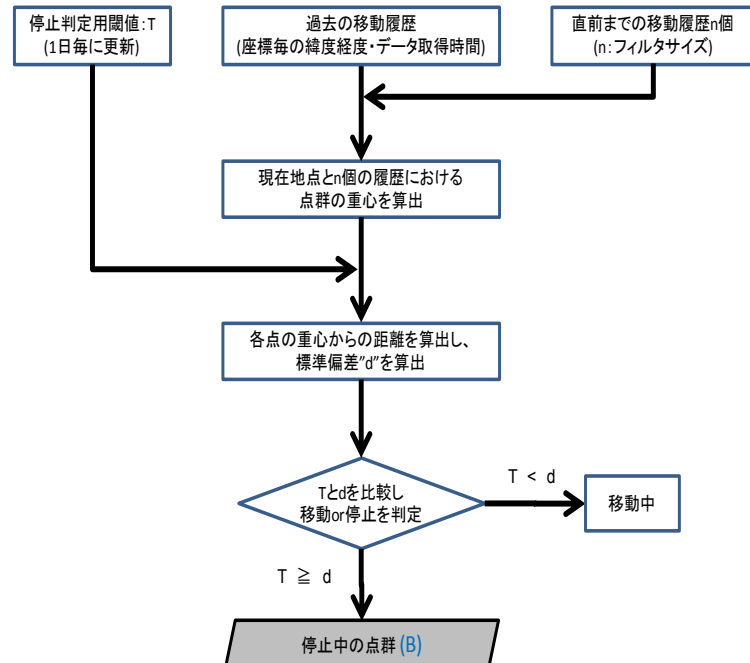


図 13 停止判定フロー

### 3.3.1.2 滞留判定

停止と見なされた点は、次に単なる一時的な停止かそれとも目的を持った滞留かの判定に進む。ここでは、移動履歴取得日当日のスケジュール情報を利用する。もし移動履歴の取得時間がスケジュールの予定時間と重複している場合、当該地点では予定を実行するために目的地に到着していると見なし、無条件で滞留中と判定する。

移動履歴とスケジュールが重複していない場合は、まず

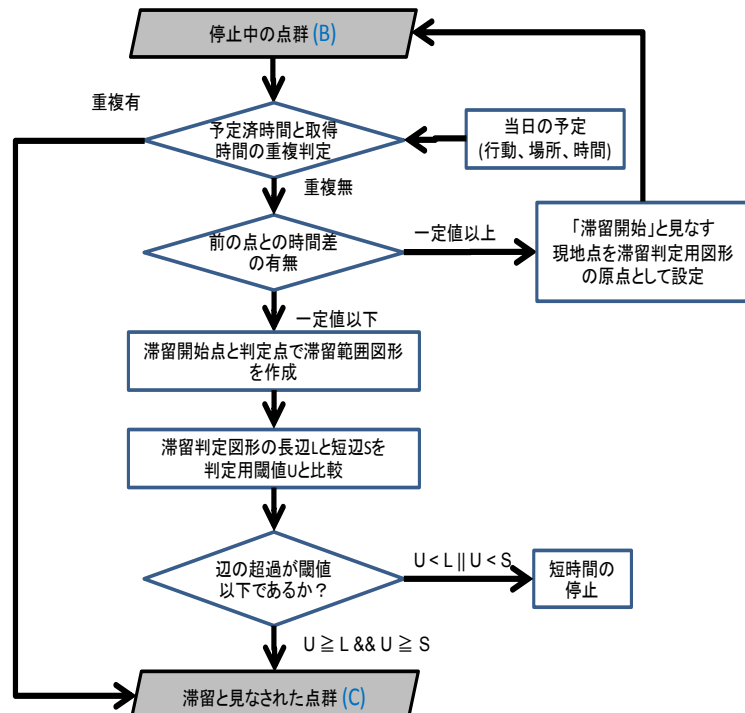


図 14 滞留判定フロー

前の滞留点との時間差を計測する。時間差が一定値以上なら当該点が滞留開始点として次の点の滞留判定へと移り、一定値以下なら前の点から滞留が続いていると見なし、連続している滞留点と当該点とで滞留範囲を形成する長方形(MBR: Minimum Bounding Rectangle)を作成する。ここで作成された長方形の二辺が図形判定用閾値  $U$  を下回った場合にのみ、当該点を滞留中の点と見なし、形成した長方形を滞留範囲として保持する。これらの計算を一日分の移動履歴全てに施し、滞留範囲を生成する。滞留範囲は、自身を構成する座標の情報(緯度・経度・時間)をそのまま有する。

#### 3.3.1.3 停止判定閾値の更新

一日分の滞留範囲を検出終わった後、生成された滞留範囲毎に重心を計算し、範囲を構成する各座標との重心距離を算出した後にそれらの標準偏差を導く。こうして生成された滞留範囲の個数分標準偏差が計算されるが、この平均を次の日の停止判定の閾値として利用する。

### 3.3.2 スケジュール内容の定量化

提案手法では、個々のスケジュールをひとつの行動と見なし、その主な構成要素である「タイトル」に着目し、個人の包括的な行動の特徴および場所毎の行動の特徴を、個人行動特徴ベクトルと場所行動特徴ベクトルとして、ベクトル空間法を用いてベクトル形式で表現する。ベクトル空間法では、特徴を多次元空間上のベクトルとして表現し、ベクトル間の類似度（コサイン相関）によって特徴の近接度合いを推し量る。

これにより個人の特徴として「どのような施設に」「どのような目的で」滞在する傾向があるかといった日々の行動や場所の特徴を明らかにできる。





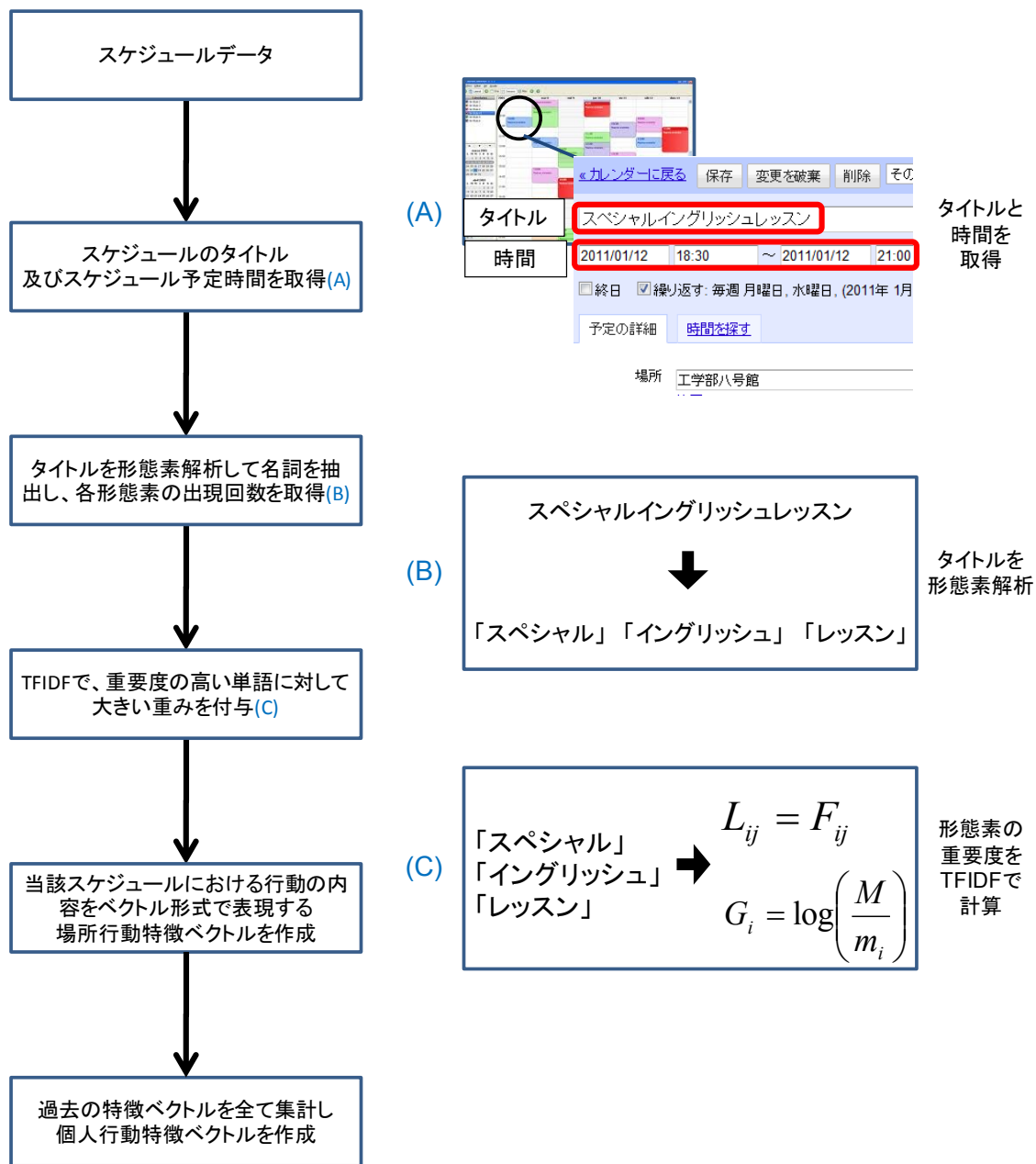


図 16 スケジュールからの特徴ベクトルの生成手順

### 3.3.2.1 個人行動特徴ベクトル

スケジュールには、ユーザの当該日時の行動内容がタイトルとして記載されている。また通常スケジュールのタイトルは行動内容を要約しており、長くても数十字程度の記述であることが多い。そのため、タイトルは名詞を中心とした語句や単語が用いられ、同一の利用者であれば表現も類似する。従って、行動の特徴を表現する要素としてタイトルに含まれる名詞に特に注目し、その出現傾向を元に特徴ベクトル(“個人行動特徴ベクトル”)を生成することで個人の行動の特徴を表現する。

行動特徴ベクトルの生成は以下の手順に従って行う。まずユーザのスケジュールに記載されているタイトルの全てに形態素解析を施す。次に形態素の出現頻度を集計し、最後に頻度に応じた重み付けを行う。こうして、全形態素分だけの次元を有するベクトルを作成する。提案手法では、この様にして生成されるベクトルを「ユーザ自身の行動の特徴」を表す行動特徴ベクトルとして用い、一人のユーザに対して一つの行動特徴ベクトルが生成される。

今ほど作成したベクトルはユーザが「“どんな” 行動を取りやすいか」を明らかにする。しかしそれが「“いつ” 行われるか」までは今のままでは窺い知る事はできない。そこで個人の行動特徴を時間帯別に計算した、“時間帯別個人行動特徴ベクトル”を別途作成する。15分間を一つの時間帯と見なし、スケジュールの開始時間と終了時間からそのスケジュールの行動がなされているであろう時間帯を判別する。その後スケジュールタイトルの形態素を解析して、出現回数をカウントする。

これにより、「環境学や経済学を勉強する機会が多い」「夜はいつも研究している」等の、ユーザが「“いつ” “どんな” 行動を取る傾向にあるか」を明確に表す事が可能となる。

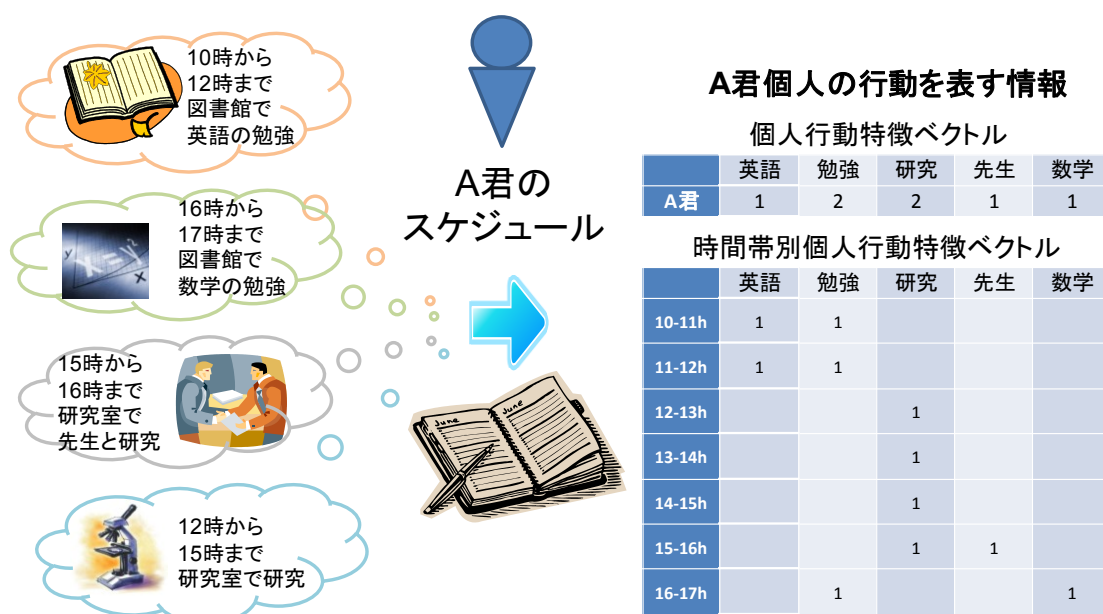


図 17 個人の行動の傾向を表すベクトル



### 3.3.2.2 場所行動特徴ベクトル

スケジュールには同時に場所の名称も記載されている。この場所は例え実質的には同じ施設を指していたとしてもユーザ毎に様々な側面を有する。例えば「柏の葉公園」という一つの場所に対して、ある人は「サッカーの試合」に出場するために赴き、またある人はその試合を「観戦」するために向かう。この例の様に、同じ場所でも人によって異なる特徴が付与されている。そこで、同一の場所を含むスケジュールを集積して場所毎に前述の行動特徴ベクトルを算出することで、個々のユーザにとってある一つの場所が有する行動の特徴を推し量ることができる。

ここで示した場所特徴ベクトルの生成手順としては、まず 3.3.1 で作成した滞留範囲において、ある一つの滞留範囲に関連付いているスケジュールをすべて集計し、それぞれ付与されているタイトルに対し形態素解析を実施する。その後は個人行動特徴ベクトルと同様の手順で、形態素の集計と重み付けを行い形態素数分の次元を有するベクトルを作成する。

また個人行動特徴ベクトルと同じく、場所特徴ベクトルでも時間帯別の行動を指し示す情報を生成する必要がある。今ほど作成したベクトルはある滞留範囲において「“どんな”行動が取られやすいか」を明らかにする。しかしそれが「“いつ”行われるか」は不明であるため、その滞留範囲での行動特徴を時間帯別に計算した、“時間帯別場所行動特徴ベクトル”が必要となる。生成方法も先の例と同様に、15分間を一つの時間帯と見なし、滞留範囲毎に付与されたスケジュールの開始時間と終了時間からそのスケジュールの行動がなされているであろう時間帯を判別する。その後スケジュールタイトルの形態素を解析して、出現回数をカウントする。

これにより、例えば「研究室では主に研究や打合せが行われる」「図書館では午前中よく勉強をしている」等の、場所に対する行動の内容がベクトルによって表され、ある滞留範囲が「“いつ” “どんな” 行動を取る場所であるか」を明確に表す事が可能となる。



図 18 場所の行動の傾向を表すベクトル

### 3.3.2.3 形態素解析

スケジュールのタイトルは自然言語でユーザに自由記述される．自然言語は名詞や動詞，助詞など数々の品詞で構成されており，特に簡潔に情報が記載されるタイトルでは名詞を中心と語句や単語が用いられ，同一の利用者であれば表現も類似する．この性質を利用して，スケジュールのタイトルから特に名詞に注目して解析することで行動の特徴の算出を試みる．本来ならば他の品詞の影響も加味すべきであるが，計算を簡素化し，スケジュールごとの表記揺れの影響を小さくするため，今回はこれらの要素を除外した．特徴の算出に適した品詞である名詞を抽出すべく，形態素解析ツール `sen`<sup>22</sup>を用いてタイトルの形態素解析を実施する．

タイトルから抽出した形態素が行動の特徴を表す要素となり，その出現傾向から行動特徴が算出される．また同様の作業をユーザの訪れた目的地に対しても行う．目的地ごとに形態素の出現頻度を調べることでユーザにとってある目的地はどのような行動のために訪れる場所であるかという，目的地ごとの行動の傾向を明らかにする．

### 3.3.2.4 TFIDF

TFIDF は，情報検索システムにおいて，重要度の高い単語を選択し高精度の検索を実現させるための手法である．ある文章中に特定の単語が頻出している場合，1)どのような文脈の文章であっても，文章を構成する上で多用される単語，2)一部の文章の中にしか頻出しない，つまりはその文章を特徴付ける単語の2種類があることが想像される．TFIDF は，特に後者の様な単語に対して高い重みを与えて特徴量を算出する手法であり，TF(Term Frequency)と IDF(Inverse Document Frequency)の2つの指標の積によって文章中の各語の重みを計算する．TF は文章  $D_j$  中における索引語  $w_i$  の出現頻度を表す．文章中に頻繁に出現する索引語に対して大きな値が与えられる．また IDF は文章集合全体にわたる索引語  $w_i$  の分布を考慮して決定される重みであり，特定の文章に集中して出現する索引語に対して大きな値が与えられる．

行動特徴ベクトルでの TFIDF は，ユーザ  $U_j$  が作成したスケジュールタイトルに出現する形態素  $K_i$  を索引語とし，ユーザ  $U_j$  の中での形態素  $K_i$  の出現頻度を TF，及び他のユーザのスケジュールにおける形態素  $K_i$  の出現分布を IDF として計算することで導く．TF を  $L_{ij}$ ，IDF を  $G_i$  とするとそれぞれ次のように導かれる．

$$\text{TF} : L_{ij} = F_{ij} \quad (1)$$

$$\text{IDF} : G_i = \log \left( \frac{M}{m_i} \right) \quad (2)$$

ここで

$F_{ij}$  : ユーザ  $U_j$  における形態素  $K_i$  の出現回数

$m_i$  : 形態素  $K_i$  を含むスケジュール数

$M$  : スケジュールの総数

(1)(2)式より, ユーザ  $U_j$  における形態素  $K_i$  の特徴量  $T_{ij}$  は以下の式より決定される.

$$T_{ij} = L_{ij} \times G_j = F_{ij} \times \log\left(\frac{M}{m_i}\right) \quad (3)$$

これを受けて, ユーザ  $U_j$  の行動特徴ベクトル  $C_j$  は, 正規化を施した上で

$$\vec{C}_j = \{T_{1j}, T_{2j}, \dots, T_{mj}\} \quad (4)$$

と表される.

また場所特徴ベクトルでの TFIDF は, ユーザ  $U_j$  が作成したスケジュールタイトルに出現する形態素  $K_i$  を索引語とし, ユーザ  $U$  が記載したある目的地  $P_j$  中でのタイトル形態素  $K_i$  の出現頻度を TF, 及び  $P_j$  以外の目的地における形態素  $K_i$  の出現分布を IDF として計算することで導く. TF を  $l_{ij}$ , IDF を  $g_i$  とするとそれぞれ次のように計算される.

$$\text{TF} : l_{ij} = f_{ij} \quad (5)$$

$$\text{IDF} : g_i = \log\left(\frac{N}{n_i}\right) \quad (6)$$

ここで

$f_{ij}$  : 目的地における形態素  $K_i$  の出現回数

$n_i$  : 形態素  $K_i$  を含む目的地数

$N$  : 目的地の総数

(4)(5)式より, 形態素  $K_i$  の目的地  $P_j$  における特徴量  $t_{ij}$  は以下の式より決定される.

$$t_{ij} = l_{ij} \times g_j = f_{ij} \times \log\left(\frac{N}{n_i}\right) \quad (7)$$

以上から, ユーザ  $U$  における目的地  $P_j$  の場所行動特徴ベクトル  $D_j$  は正規化を施した上で

$$\vec{D}_j = \{t_{1j}, t_{2j}, \dots, t_{mj}\} \quad (8)$$

と表される.

### 3.3.2.5 定量化済みスケジュール情報の作成

スケジュールのタイトルを解析し場所行動特徴ベクトルを作成した後, 元のスケジュール情報に対し作成したベクトル情報を追加する. これにより最終的なスケジュール情報は

・タイトル

- ・場所名
- ・予定開始時刻
- ・予定終了時刻
- ・場所行動特徴ベクトル

の 5 種類のデータからなる情報となる。

### 3.3.3 滞留範囲とスケジュール内容の統合

滞留範囲にはその範囲を構成する座標が複数個保持されているが、各座標には座標がデバイスによって取得された時間が付与されている。またスケジュール内容にも同様に時間情報が含まれている。これら両者が有する時間情報において、滞留していた時間とスケジュール上で登録されていた予定時間の間に重複が見られる場合、このスケジュールは当該滞留範囲で実行されたものであると見なし、当該滞留範囲にスケジュール情報を統合する。これを全ての滞留範囲及びスケジュール情報に対して実施し、この作業の完遂をもって予測用データセットが完成する。

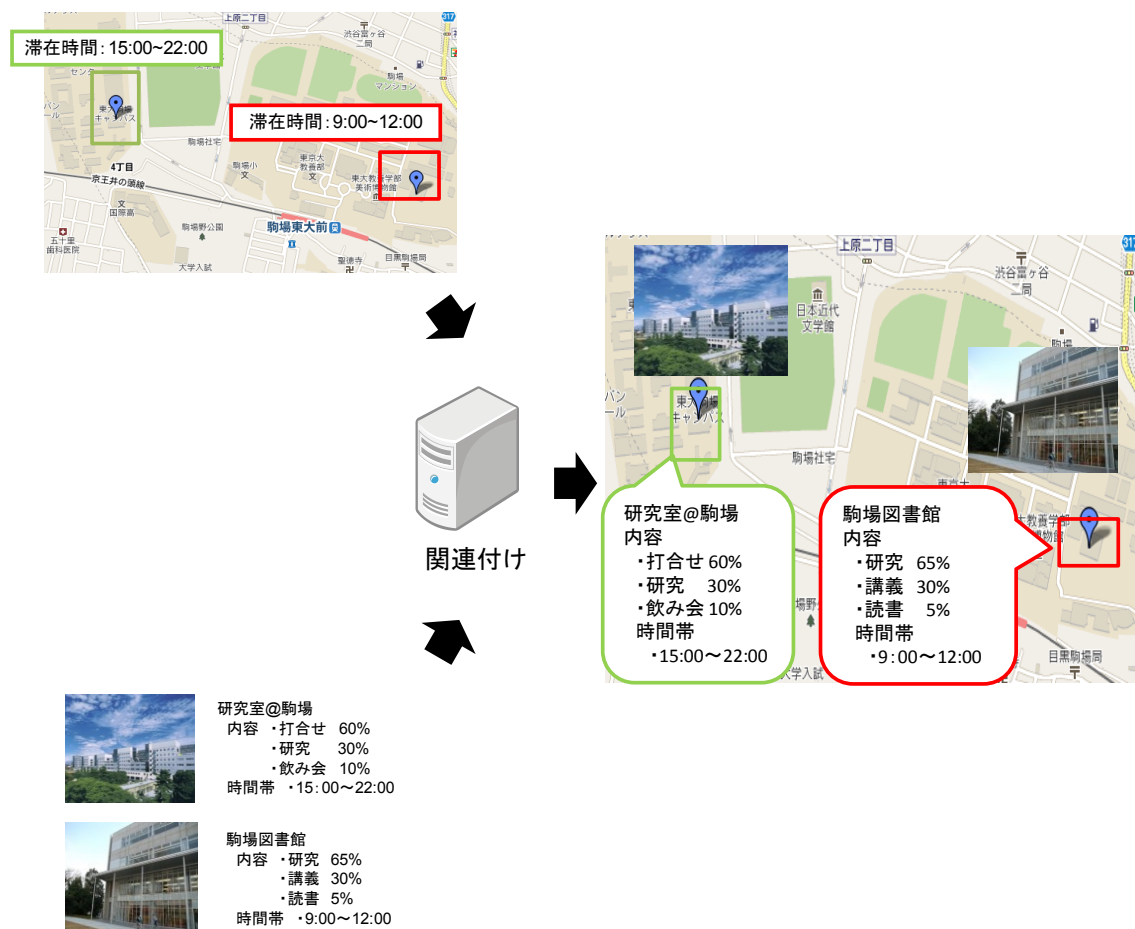


図 19 滞留範囲とスケジュール内容の統合

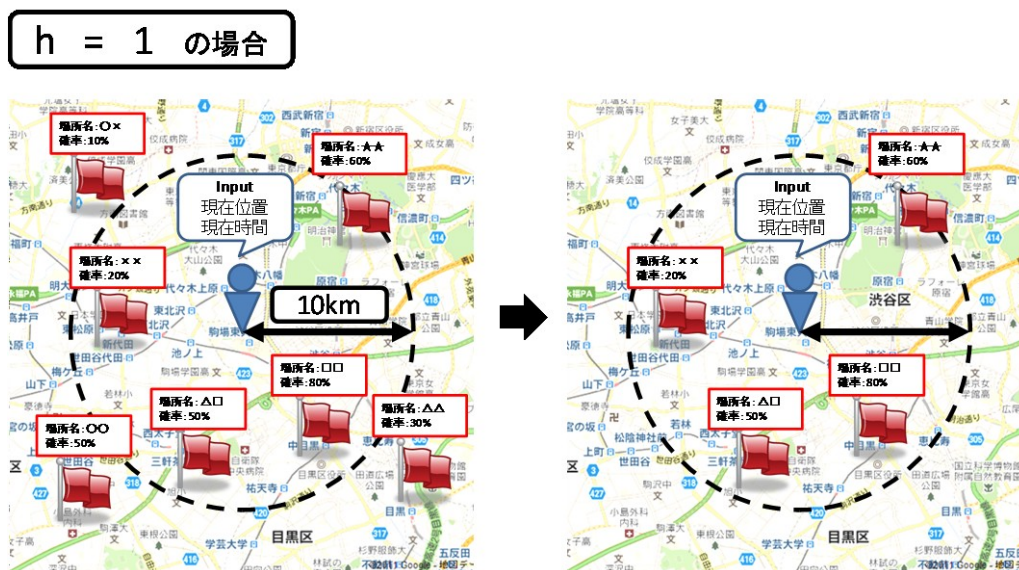
### 3.4 目的地の予測

完成した予測用データセットを用いて、ユーザが訪れようとする目的地の予測を行う。予測は図 21 に示す手順に沿って行われる。

まず、ユーザの現在位置を GPS 受信端末により取得する(①)。この情報には位置の取得時間も含まれるため、自動的に現在時間も判明する。

次に現在地点から一定の時間範囲を設定し、その時間範囲で到達可能な目的地のみを予測する。例えば、「一時間後に到着可能な目的地の候補」「三時間後に到着可能な目的地の候補」のように、現在時刻と予測される目的地での到達時刻との時間範囲を任意に設定する(②)。ここでは時間範囲を  $h$  時間とし、3.3 で作成した予測用データセットの中から、 $h$  時間以内に物理的に到達可能な滞留範囲を検索・抽出する(③)。ここで「到達可能」とは、ユーザの現在地点から滞留範囲の中心地点まで直線距離で、時速 10km で  $h$  時間以内に移動可能である状態を指す。これは都市部における移動、即ち徒歩+公共交通機関の使用かつ煩雑な道路網通過する場合の移動を想定し、直線距離に換算した際の実質的な移動速度と仮定したものである。現在位置とデータセット内全ての滞留範囲までの距離を計算し、現在位置から半径  $h \times 10\text{km}$  以内に存在する滞在範囲を現時点での目的地候補と見なす(④)。

次いで、滞留範囲に付与されている場所行動特徴ベクトルと、ユーザの個人行動特徴ベクトルの類似度を算出する(⑤)。これにより、「ユーザの行動の傾向と、ある場所で日頃行われている行動の傾向」がどの程度似通っているかが計算可能となる。さらに、時間帯別場所行動特徴ベクトルと時間帯別個人行動特徴ベクトルの類似度を算出し(⑥)、「目的地到着予定時間にユーザが日常的に行う行動と、同時刻にある場所で行われている行動の傾向」の類似性を算出する。最後に両方の類似度を足し合わせ、その結果が最も大きい値を示した滞留範囲を、現時点で最も訪れる可能性の高い予測目的地と見なす(⑦)。



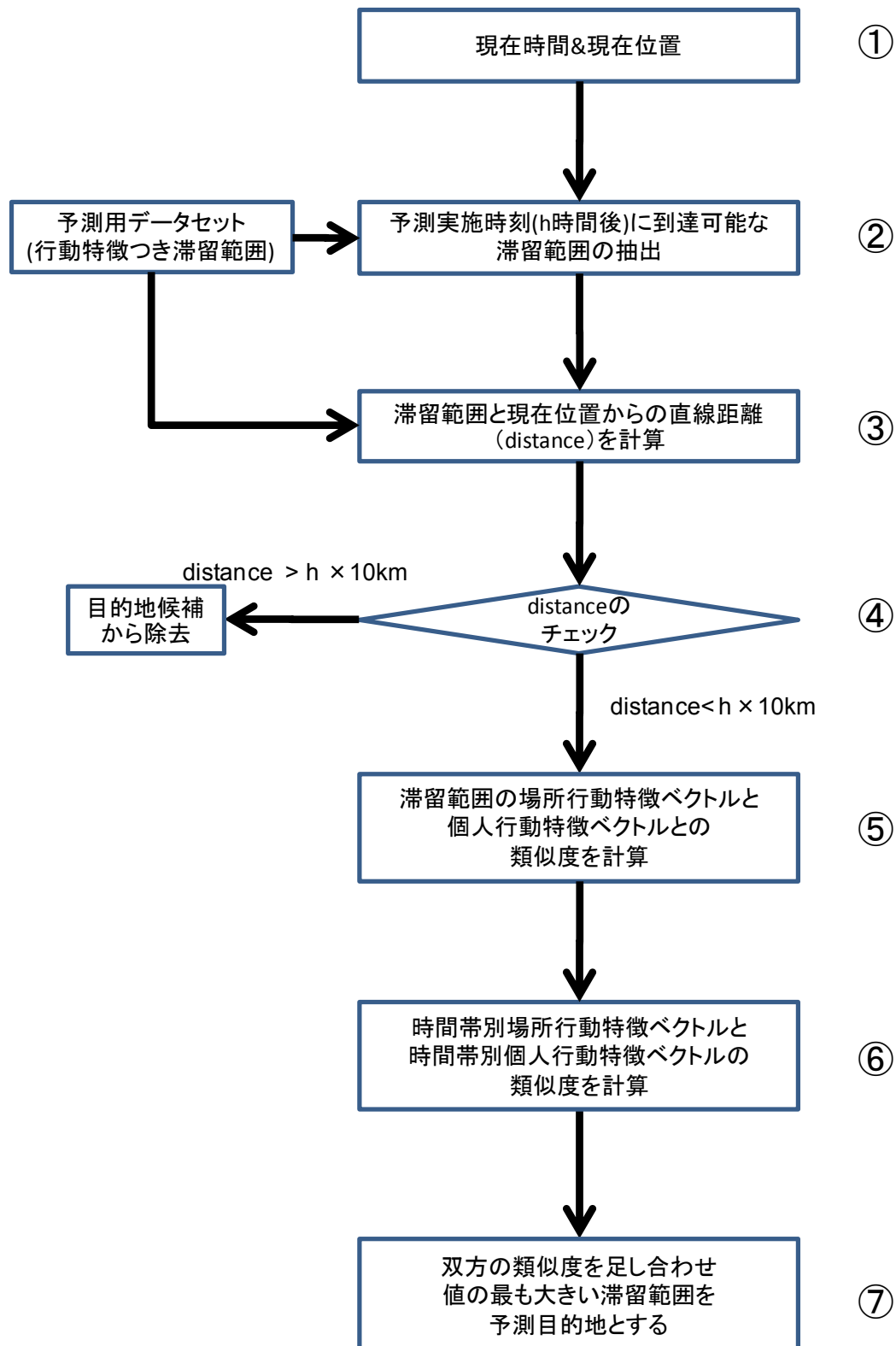


図 21 目的地予測の計算過程

## 4. 提案手法の実証と評価

前章で示した提案手法の性能を評価するために、手法のプロトタイプを Java アプリケーションによって実装し、評価実験を実施した。

目的地予測の精度については、次の観点での評価を行った。

- 目的地予測に必要な予測用データセット作成用の学習期間に対する精度評価
- 一日の時間帯別の予測精度評価

実証に際して、被験者 1 名の位置情報とスケジュール情報を 1 分間隔で三ヶ月分収集した。また同様に被験者 3 名の個々の位置情報とスケジュール情報を一ヶ月分収集した。以降の実証では三ヶ月分のデータを用いて各観点からの精度評価を行ない、5 名の一ヶ月分のデータからは異なるユーザ間における予測精度の発現具合を検証した。実証開始時には各々のデータにおいて二週間のデータセット作成用期間を設け、以降のユーザの移動に対して一時間おきに目的地予測を実施した。また、予測が実施される毎にデータセットは更新され、ユーザの新たな行動特徴を計算に反映させる。実施後は、予測結果と実際のユーザの行動を記録した実績値とを比較し、予測精度の算出を試みることにした。

予測精度は、現在時刻から一時間後に訪れると予測された結果(以後、一時間後予測と呼称)と実際の目的地とが一致した回数を予測回数で除することで求める事とした。またこれとは別に、三時間後・五時間後に訪れると予測された結果(それぞれ以後三時間後予測、五時間後予測と呼称)についても同様の計算で予測精度を求めることにした。

図 22 に、三ヶ月分のデータによる目的地予測実施結果及び各期間においてユーザが記録したスケジュールの個数を示す。これを見ると、予測開始時点では 52.9%であった一時間後予測の精度が、日々データセットが更新される毎に精度を増し、30 日後には最高値の 72.6%、最終的に予測期間終了時には 72.4%の予測精度を記録した。また三時間後予測と五時間後予測に関しては、それぞれ予測開始時点での予測精度が 50.0%、46.7%であったものが、予測期間終了時には 64.9%、53.4%の精度を示した。また図 23 には別の被験者 3 名の行動履歴データを用いた目的地予測結果を表す。ここでも、一時間後予測の最終的な予測精度は 3 名の結果の平均で 70.3%を示している。行動の傾向が異なる複数の被験者から収集したデータにおいて恒常的に高い予測精度が表れたことで、提案手法の適応性の高さを実証できたと言える。

図 22 を詳しく見ると、10 日～20 日程度で予測精度の安定が見られる。これはユーザの日常的な行動に関しては当該期間内に網羅が完了し、ユーザの行動傾向及び各々の場所で行われる行動の傾向の大部分を把握できたことで、ユーザがこれから取るであろう行動を的確に判断でき、またその行動が行われるであろう場所を高い確度をもって予測できるようになったからであると考えられる。

また、予測開始後 40 日に関しては若干予測精度が低下している。これに関しては、当該期間において、日常的な行動の傾向から外れたイレギュラーな移動が多数行われたためで



ある。図 22 から 40 日時点のスケジュールの数を見てみると、過去に既出している日常的な予定が二個記載されていたのに対し、これまで行った経験のない新規の予定が五個記載されていた。この時点で、ユーザの行動の傾向を表す行動特徴ベクトル内には新たな行動を表す要素が含まれていなかったため、行動特徴に沿った予測が上手く働かず、結果予測精度の低下を招いたと考えられる。

図 24 は、時間帯毎に見た予測精度の変遷である。ここでは、一日を一時間毎に分割し、それぞれの時間範囲において目的地予測の精度を計算した。ここで表す予測精度は、各々の時間範囲における一時間予測・三時間予測・五時間予測の三種の結果を平均したものをを用いている。またユーザの移動が発生せず予測が行われなかった 0 時～7 時及び 23 時～翌 0 時までについては、図表への記載を割愛した。

これを見ると、一日の活動時間帯、主に 9 時～16 時における予測精度は 6 割を超え、中には 8 割に近い精度を示す時間帯も表れている。これは業務や講義といった日々のルーチンワークは当該時間帯に実施されることが多く、しかも行動の内容も同一で有る場合が多いため、目的地予測が容易になるためであると推測される。

それに比べ、17 時～22 時における予測精度は軒並み 6 割を下回っている。業務時間が終了する時間帯である 17 時以降は、ユーザの嗜好や興味によって行動が左右される割合が高く、決まった場所で決まった行動を取る事が業務時間内に比べ少なくなる。そのため行動の特徴を掴みづらく、それに伴い予測精度の向上が妨げられてしまったと考えられる。

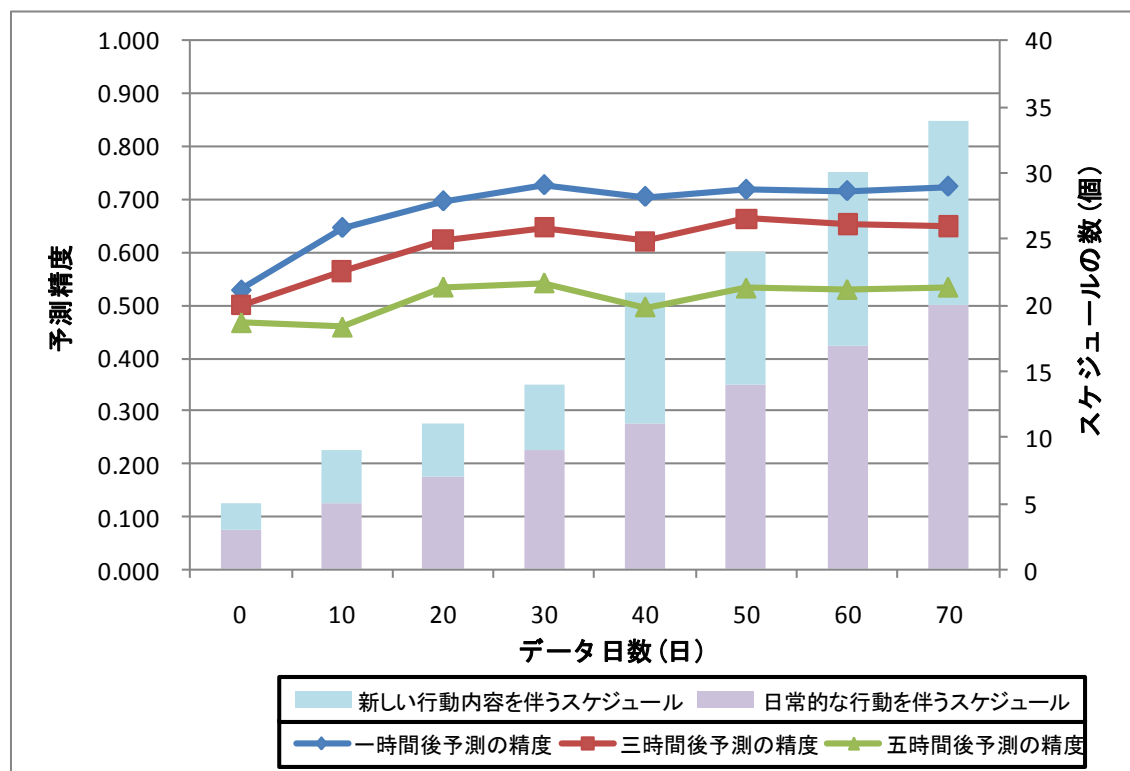


図 22 経過期間に対する目的地予測精度の推移



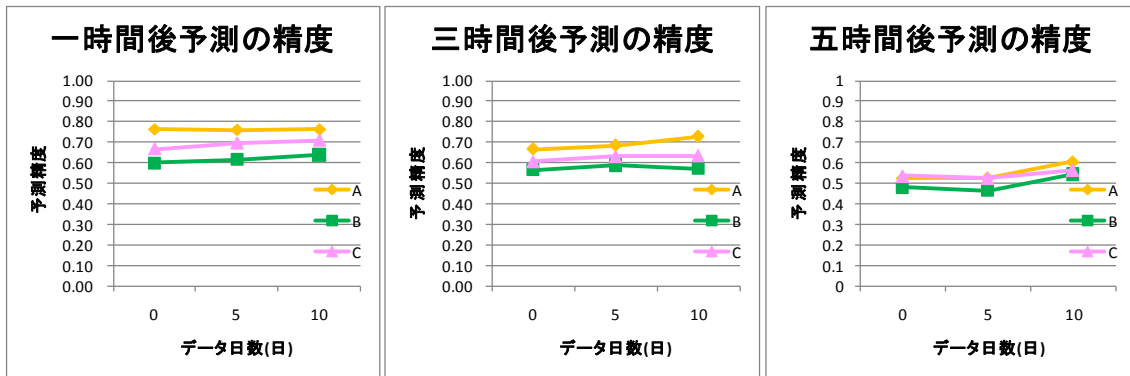


図 23 被験者 A, B, C における, 経過期間に対する目的地予測精度の推移

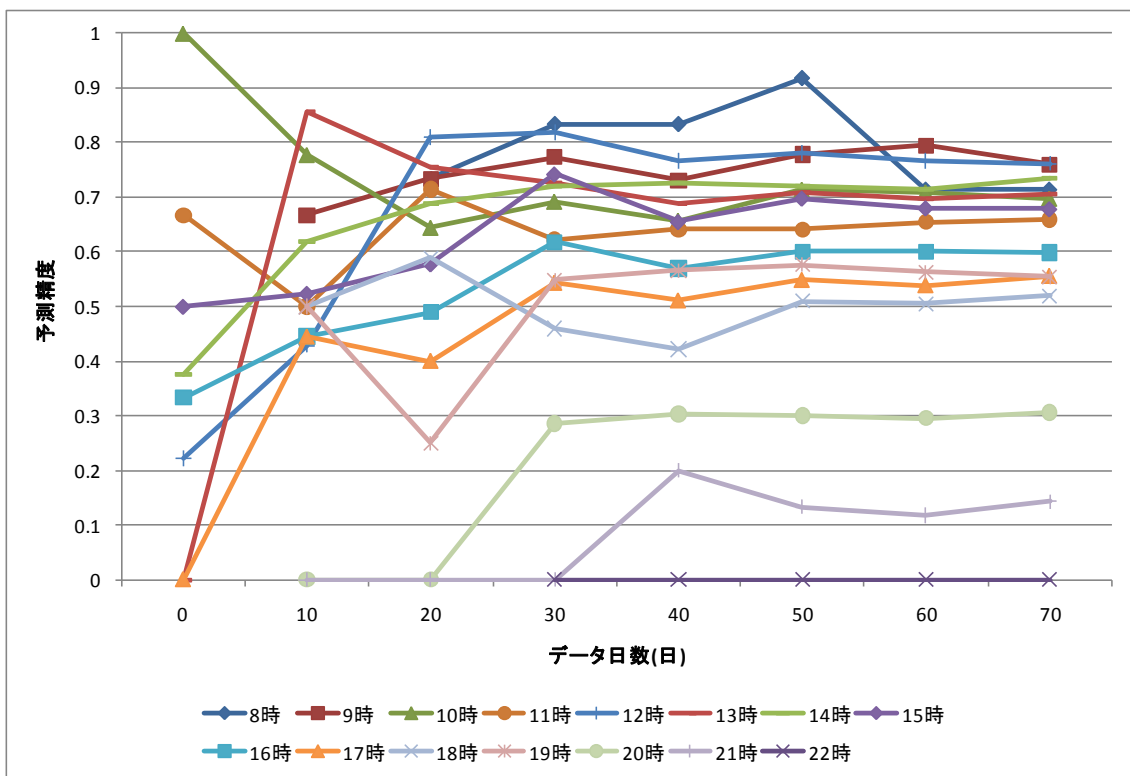


図 24 学習期間に対する時間帯別目的地予測精度の推移

## 5. 結論と今後の課題

本研究では、今後の位置情報活用社会において、不可欠な存在たり得る“目的地予測”について、数値的な位置情報をスケジュールに基づいたテキスト情報で補完し、予測に用いるというこれまでに無い視点からのアプローチにより、従来手法よりも高精度な予測手法を構築することを目指した。

提案手法を実証した結果、被験者数が十分とは言えないが、従来の目的地予測手法よりも高い精度での目的地予測が可能であることを示した。

提案手法を更に発展させ、より高精度な目的地予測手法を開発・構築するための課題として、第一に状況別行動特徴ベクトルの作成及び予測処理への組み込みが挙げられる。提案手法では、ユーザの行動の内容を表す特徴ベクトルとして「個人行動特徴ベクトル」「時間帯別個人行動特徴ベクトル」の二つ、またある場所における行動の内容を表す特徴ベクトルとして「場所行動特徴ベクトル」「時間帯別場所行動特徴ベクトル」の二つをそれぞれ計算し予測に用いた。しかし、これだけの情報では、ユーザの行動の特徴や場所における行動の特徴を完全に把握することは難しい。即ち、これ以外にも例えば男女間での行動特徴の差異、平日と休日の行動の違い、夏と冬等季節毎の予定の変化といった、日々のユーザの行動特徴に変化をもたらす要因は多く存在している。従って、全体的な行動の特徴と時間帯毎の行動の特徴のみで目的地を予測するよりも、よりユーザの行動内容の変化に対応した特徴ベクトルを作成した上で、その行動特徴に見合った目的地を予測することが可能になれば、更なる予測精度向上が期待できる。

また第二の課題として、行動特徴ベクトルの作成に際して時間経過に伴う行動特徴の変化を考慮する必要がある。本研究の提案手法で行動特徴を計算する際は、過去の古いスケジュールと現在のスケジュールを特に区別せず、全て同列として行動特徴ベクトルを算出している。しかし、行動の傾向は時を経るにつれて大きく変化や、前述の季節による変容が想定される。例えば‘スポーツクラブに通い始めた’‘就職して仕事をするようになった’等、現在と過去の行動特徴には時間経過に伴って差異が発生することが予想され、その場合は現在の行動の傾向に沿った場所が最尤な目的地として予測されるべきである。従って、過去よりも現在の行動により高い重みをつけて行動特徴を計算することで、常にユーザの最新の行動特徴に適応した目的地予測を実施することが可能となる。

さらに、行動特徴及び場所特徴にユーザの行動の傾向をより正確に反映させる事も必須である。ユーザの取る行動の傾向と、ある場所で行われる行動の傾向をより正確に取得する事は、目的地予測の精度に直結する。そのためには、行動の内容を特徴付けるデータとして使用するテキスト情報の多様化が不可欠である。本研究では、ある場所での行動を特徴付ける要素として、スケジュール情報を用いた。これはスケジュールには“スケジュールのタイトル”“場所”“時間”の情報が基本的に全て記されており、行動の特徴を特定し易くかつ同じ時間情報を有する位置情報との統合が容易になるためであった。その反面、

日常の些細な行動，例えば「帰宅途中にデパートで買い物」「通勤中に喫茶店で朝食」等といった予定はスケジュールには記載されないことが一般的であり，こうした細かい行動も把握し目的地予測に取り入れることで，よりユーザの行動を詳しく予測することが可能となり，ユーザごとにカスタマイズされたサービスの提供が可能となる．スケジュール以外にも行動の情報を取得できる情報は多々ある．例えば店舗での購買に付随するレシートからは，店舗名と購入した品目，購入した時間を読み取ることでユーザの訪れた詳細な位置と時間，またユーザの嗜好が得られる．さらに，ユーザが日々つけるブログもその一つである．ブログには“その日の行動”“行動を行った場所の名称”“行動を行った時間”が事細かに記載されていることが多々ある．ブログに含まれる情報を解析することで，ユーザの行動の特徴をより詳細に算出することが可能になると思われる．また近年脚光を浴びている **Twitter** も行動特徴の抽出に大きく役立つと考えられる．**Twitter** ではユーザが現時点で行っている行動及び現在地を「ツイート(つぶやき)」という形で記録・世間に公表している．また，**GPS** 機能付き携帯端末で **Twitter** を利用しているユーザであれば，**GPS** によって取得した位置情報もツイートに添付して記録できる．加えてツイートの時間は全て履歴として残されている．これらユーザの行動を特徴付ける情報をあらゆる角度から取り入れ解析することで，さらなる目的地精度向上をはかる必要がある．

## 謝辞

本研究の進行に当たっては，指導教官である柴崎亮介教授に終始多大なる御指導御鞭撻を賜りました．また副査の瀬崎薫教授と丸山祐造准教授からは，提案手法内の問題点や今後の課題について数多くの御指摘を頂きました．その他にも，研究の核となる目的地予測手法の実装に際し無知な私を一から御指南頂き，研究の内容に関する的確な助言を下さいました空間情報科学研究センターの金杉洋特任研究員，そして手法の実証に欠かすことのできない各種実測データの収集に御協力頂きました新領域創成科学研究科所属の学生有志の皆様に，心より御礼申し上げます．

## 参考文献

- 1) 総務省,情報通信統計データベース : 携帯電話・PHS の加入契約者の推移,  
<http://www.soumu.go.jp/johotsusintokei/new/index.html>, 平成 22 年 12 月 12 日
- 2) 総務省, : 事業用電気通信設備規則,平成 18 年 1 月 5 日総務省令第一号
- 3) NTT ドコモ, : i コンシェル,<http://www.nttdocomo.co.jp/service/customize/iconcier/>
- 4) ケータイ Watch, : i コンシェルが情報を繋ぐ, ドコモが描く携帯のビジョン  
[http://k-tai.impress.co.jp/docs/event/interop2010/20100611\\_373742.html](http://k-tai.impress.co.jp/docs/event/interop2010/20100611_373742.html), 2010 年 6 月 11 日
- 5) NAVITIME JAPAN, <http://corporate.navitime.co.jp/index.html>
- 6) Twitter, : <http://twitter.com/about/>
- 7) foursquare, : <http://foursquare.com/about/>
- 8) 株式会社コロプラ, : <http://colopl.co.jp/>
- 9) 株式会社マピオン, : ケータイ国盗り合戦, <http://kntr.jp/pc/>
- 10) 株式会社マピオン, : ニュースリリース「ケータイ国盗り合戦, Android スマートフォン向けアプリの提供を開始」, [http://blog.mapion.co.jp/release/2010/12/101224\\_14846.html](http://blog.mapion.co.jp/release/2010/12/101224_14846.html)  
2010 年 12 月 24 日
- 11) 一般社団法人モバイルコンテンツフォーラム, : モバイルコンテンツ関連市場に関する市場規模調査,2010
- 12) 大西啓介, : 「NAVITIME」の世界戦略と位置情報ビジネスの可能性, ITU ジャーナル Vol.39 No.11, 2009.11
- 13) 株式会社コロプラ, : ニュースリリース「位置ゲーのコロプラ, ユーザー数が 100 万人を突破」, [http://colopl.co.jp/press/news/2010/0402\\_1.html](http://colopl.co.jp/press/news/2010/0402_1.html), 2010 年 4 月 2 日
- 14) Chitika Inc., <http://chitika.com/>
- 15) Cirius Technologies Inc, <http://www.cirius.co.jp/corporate/>
- 16) D Ashbrook, T Starner, : Using GPS to learn significant locations and predict movement, Personal and Ubiquitous Computing Vol.7 Issue.5, pp.275-286, 2003
- 17) 西野正彬,山田智広他 : 時刻情報を含む特徴滞在パターンを用いた行動予測方式の検討, 2009 年電子情報通信学会総合大会 情報・システム講演論文集, 2009
- 18) 山田直治,他 : GPS 搭載携帯電話を用いた移動経路履歴に基づく訪問地・経由地予測システム, 情報処理学会研究報告 110(130), pp.47-54, 2010.
- 19) C. Zhou, S. Shekar, and L. Treveen: Discovering Personal Paths from Sparse GPS Traces, JCIS 2005 Workshop on Data Mining, 2005
- 20) Juhong Liu, Ouri Wolfson, Huabei Yin, : Extracting Semantic Location from Outdoor Positioning Systems, Proceedings of the 7th International Conference on Mobile Data

---

Management,IEEE,2006

21) iCal4j, : <http://ical4j.sourceforge.net/introduction.html>

22) 形態素解析ライブラリ Sen, : <https://sen.dev.java.net/>