サポートベクターマシンによる GPSデータの移動/滞在の自動判別手法

水野 敬太[1]・金森 亮[2]・佐野 渉二[3]・中島 秀之[4]・伊藤 孝行[5]

[1]非会員　名古屋工業大学大学院　工学研究科情報工学専攻（〒466-8555 名古屋市昭和区御器所町）
E-mail: mizuno.keita@itolab.nitech.ac.jp

[2]正会員　名古屋工業大学しくみ領域（〒466-8555 名古屋市昭和区御器所町）
E-mail: kanamori.ryo@nitech.ac.jp
[3]非会員　はこだて未来大学（〒041-8655　北海道函館市亀田中野町116番地2）
E-mail: sano@fun.ac.jp
[4]非会員　はこだて未来大学（〒041-8655　北海道函館市亀田中野町116番地2）
E-mail: nakashima@fun.ac.jp
[5]非会員　名古屋工業大学大学院 産業戦略工学専攻／情報工学科（〒466-8555 名古屋市昭和区御器所町）
E-mail: ito.takayuki@nitech.ac.jp

本研究では，近年普及しているGPS携帯端末を利用して位置情報を収集し，サポートベクターマシンを用いたリアルタイム処理でGPSデータを移動/滞在に自動判別手法を提案する．函館市にて実施されたGPS携帯端末を利用した交通行動調査データを利用し， サポートベクターマシンによる判別精度が高いことを確認した．また交通需要予測モデルの構築として，簡単な交通手段選択モデルのパラメータ推定を行い，推定に結果について考察をおこなった．

Key Words : *Smart Access Vehicle System, Support Vector Machine, GPS*

1. はじめに

本研究では，スマートシティはこだてプロジェクトで実施されているGPS携帯端末を利用したパーソントリップ調査データの基礎分析として，GPSデータから移動/滞在をリアルタイム処理で自動判別する手法の検討を行う．スマートシティはこだてプロジェクトでは，はこだて圏を対象として，街の様々な活動やサービスを有機的に統合することで，全体として住みやすい便利な街を目指す．街の生活者や観光客に対して，外食，医療，および観光等を連携させた様々なサービスを創出することで，より活性化された街を実現できると考えている．一方，様々なサービスが連携した街を実現するためには，広範囲に分布しているそれぞれの施設やスポットを利用者にストレスを感じさせることなく，短時間に高効率で結ぶ交通システムを整備する必要がある．施設を高効率で結ぶ交通システムのため，都市内のすべてのバスやタクシー等の公共交通機関の運行を集中管理する高効率の交通システムであるSmartAccessVehicleSystem（SAVS）を実現することを目指している．

SAVSを実現するための中心となる考え方は，固定路線や固定ダイヤを全く持たずに，利用者の需要に応じてバスの路線や運行ダイヤを決める運行方式であるフルデマンドバスシステム[1]である．我々は，SAVSにおいて，フルデマンドバス方式をすべての公共交通機関に適用することを考えている．SAVSでは，利用者は現在位置，目的地，および時間を運行主体に連絡して，移動をリクエストすることで，すべての車両の現在位置と運行予定ルートを把握している運行主体は，車両の中から利用者の呼び出しに最適の1台をリアルタイムに選択して，選択された車両のルートを調整し，利用者に乗車予定時刻と目的地への到着予想時刻を告げる．車両が発車後にも他の利用者のデマンドが発生することを考慮し，目的地への到着予想時刻には，時間の幅を設ける必要がある．乗車予定時刻と目的地への到着予想時刻を利用者が受け入れれば契約が成立し，選択された車両の運行ルートを変更する．フルデマンドバスシステムに対して，単純なモデルで行われたシミュレーション結果によると，少数台の車両で運用する場合には利用者のデマンドに応じて寄り道する確率が高くなるため，固定路線の場合に比べ

非効率であるが，車両台数を増やすに従って効率が良くなり，ある程度以上の車両台数になると路線バスより所要時間が短くなることが示されている[2]．しかし，車両台数をどの程度増やせばよいのかという具体的な数値は分かっていない．そこで，はこだて圏におけると生活者と観光客や道路網に対して，SAVSが有効であるかを検証するため，また，SAVSの効率や乗車料金，運用コストを考慮しながら車両台数を決定するために交通需要予測にNestedLogitモデルを用いた人流シミュレータ[3]を構築予定である．デマンドバスの運行シミュレータに用いるデータとして，2012年12月からGPSを内蔵したスマートフォンを用いた交通行動調査[4]を行っている．

本論文では，函館でおこなわれた交通行動調査で収集されるデータから交通需要予測モデルを作成するために，Support Vector Machine（SVM）を用いてGPSデータの滞在と移動の自動判別手法を提案する．函館でおこなわれた交通行動調査では，被験者の移動のGPSデータが収集されているが，一般的な実際に滞在した場所をユーザーは確認しないため（高齢者の行動データ取集のため，調査負担の軽減から確認作業をなくし，なるべくGPS携帯端末を持ち歩くだけにしている），何らかの方法で特定する必要がある．判別には，時間的に連続したGPSデータ群の中のあるGPSデータの前後のGPSデータから速度や距離の情報を作成し，利用した．滞在地点を抽出することにより被験者がある地点から別のある地点までどのくらいの時間をかけて移動したのかという情報が得られる．また，交通需要予測モデルの作成には徒歩，自動車，バス，および電車の移動などの移動手段の情報も必要になるが，本論文では，移動手段の自動判別はしておらず，スマートフォンを利用したアンケート調査結果（移動開始前の行動意図の入力結果）から移動手段の情報を得ている．自動判別手法としてはSVMを適用し，高い判別精度が得られることを確認した．

またSVMにて自動判別されたトリップ情報とアンケート調査結果から得られる移動手段から簡単な交通需要予測モデルの構築として，交通手段選択モデルのパラメータ推定を行った．

2.　函館交通行動調査

函館で実施された交通行動調査の収集の方法，収集される情報の形式，および収集された情報の詳細についての説明を述べる．

(1)　スマートフォンアプリケーションを用いた交通行動調査

函館で実施された交通行動調査では，スマートフォンのアプリケーションを作成し，被験者の移動に関する情報を収集した．アプリケーションを利用して，収集することのできるデータは，被験者の移動予定，実際に移動した際のGPSの移動軌跡，および被験者が録音した音声データである．

移動予定は，被験者に一日の移動予定について，移動の目的，目的地，および交通手段を入力してもらう．それぞれの入力は選択式になっており，表1で示す選択肢の中から当てはまるものを選択する．

表 1 移動予定の項目と選択肢

項目	選択肢
目的	帰宅，通勤・通学，飲食，買い物，遊び，通院，ドライブ，その他
目的地	家，会社，学校，飲食店，スーパー，娯楽施設，その他のお店，商店街，公園，病院，その他
交通手段	徒歩，自転車，自動車，タクシー，バス，電車，路面電車，飛行機，船

位置情報の計測はアプリケーションの開始ボタンを押すことでおよそ30秒に1回の周期でGPSによる位置情報計測をおこなう．

音声録音機能は，被験者がアプリケーションの録音ボタンを押すことで，被験者の声を録音することができる．

作成したアプリケーションを用いて，実行された交通行動を取得するだけでなく，あらかじめ予定を入力し，入力した予定が変更時に予定の修正や理由を入力することで，人々はどのような理由で予定を変更するのかという情報を得ることも可能になっている．すなわち，家からの出発前には，一日の予定として，行動目的，目的地，および交通手段を決定している範囲で入力し，移動中に予定に変更があるときは，変更以降の予定をすべて入力することで予定の修正をおこなった上で，変更した理由を録音することで交通機関の選択理由を取得できるようになっている．

(2)　交通行動調査の実施

冬の調査として，2012年12月下旬から2013年4月下旬までの4ヶ月間，交通行動調査アプリケーションを使用して，交通行動調査をおこなった．位置情報を計測するために被験者の行動がわかるために，プライバシー保護の観点から被験者の匿名性を高めるために，特定非営利活動法人スマートシティはこだてにアプリケーションをインストールしたスマートフォン20台を預け，被験者として20代から70代までの男女計20名の選出と交通行動調査の運用を委託した．調査開始時には，性別，年齢，運転免許の有無，および運転頻度などのアンケートも依頼

した．

(3) 交通行動調査の結果

2012年12月下旬から2013年2月下旬までの約2ヶ月分のデータを暫定的に入手した際の調査結果について示す．

図1は函館交通行動調査で収集したGPSデータの軌跡を白い線でGoogle Earth上に描画したものである

図 1 函館交通行動調査で収集された GPS データの軌跡を描画した図

図1からわかるように函館交通行動調査で収集されたデータは，函館全域をカバーしている．図1上で図の外に軌跡が飛んでいるデータは，GPSの観測エラーデータであり，モデル作成などの際には取り除いている．

本調査のように，同一人物の長期間に渡る調査をおこなう調査はみられないので，本調査においてどの程度のデータが取得できているかについて述べる．比較対象は，1999年に函館圏(函館市・上磯町・大野町・七飯町)で行われた都市交通特性調査(パーソントリップ調査:PT調査)[5]を用いる．PT調査[6]では，出発地点から目的地点まで「だれが」「いつ」「なぜ」「どこへ」「何を使って」移動するかについて調査することを目的とし，アンケート形式で全国規模で行われており，調査結果は都市交通計画や防災，福祉などさまざまな分野で活用されている．我々が行う交通行動調査では，アンケート形式で行われている従来のPT調査で収集されているものよりも詳細なデータ，すなわち，位置情報計測による移動軌跡や予定変更時の理由なども収集して人々の行動の特性をより詳しく把握することを目的としている．また，データを収集する年月も違うため，必ずしも同じようなデータを収集しようとするわけではないが， 1つの指標として，函館圏で行われたPT調査の結果を用いる．

本論文では，函館圏におけるPT調査と比較を行うために，ある目的による出発地から目的地までの移動を1トリップとしたトリップ数の結果について述べる．

表 2 本交通行動調査と過去の PT 調査の比較

	本交通行動調査	PT調査
調査年月	2012年12月~2013年2月	1999年
調査人数[人]	20	14400(総人口36万，抽出率4%)
自動車保有率[台/100人]	85.0	56.8

本調査では，外出時には予定の入力を依頼しているため，予定の入力があった日を外出した日と判断する．つまり，外出していても入力漏れがある場合は，外出していないものとする．また，連続して同様の予定が入力されている場合は予定の修正を行ったものと判断し，前に入力されてあったものはトリップ数にカウントしない．

本調査と函館圏におけるPT調査では，外出率はほぼ同じであり，外出する日のみの平均トリップ数のすべての日に対する増加率も，本調査では1．27倍，PT調査では1．33倍と同程度であったが，平均トリップ数は，本調査の方が少なかった．

本調査において，平均トリップ数が少なくなった原因としては，まず，被験者の構成が実際の人口分布，属性分布とは異なることが考えられる．つまり，本調査では，世代別や職業別，家族形態別などさまざまな方の行動特性を詳しく把握する目的から，被験者として20代から70代までの男女が均等になり，さまざまな職業や家族形態の方が選ばれるように依頼したため，実際の人口分布，属性分布とは異なることによる影響が出ていることが考えられる．2つ目の理由としては，データの入力漏れや現在のアプリにおける目的，目的地の項目の不備があると考えられる．データ漏れについては，スマートフォンの携帯やスマートフォンの入力を忘れたこと以外にも，「帰宅」の入力をされていない場合がよく見られた．具体的には，会社勤めの方の勤務日については「通勤」と「帰宅」の両方が入力されていたが，休日に家から本屋に行き，帰宅する場合などで，「帰宅」の入力がされていないことがあった．また，家から本屋とアパレルショップに行き，帰宅する場合，実際は3トリップであるが，作成したアプリで入力される項目としては，本屋へ行くのもアパルショップへ行くのも，「買い物」，「その他のお店」の組み合わせであるため，家から本屋とアパレルショップに行くことがまとめられ， 2トリップの入力しかないことがあったと考えられる．本稿では，被験者によるスマートフォンの入力がどの程度行われたかを述べるために，明らかに「帰宅」の入力がない場合でも「帰宅」の項目を加えるようなデータ補完はしていない．

本調査においても，従来のパーソントリップ調査に近

いトリップ数のデータが得られると考えられる．今回の調査では，長期間行うこと，スマートフォンの操作に不慣れな人もいることを考慮し，できるだけ操作を簡単にしたアプリを作成したため，朝に入力した予定を夜に確認することや，予定の変更をする際に変更分だけを入力することはできない．

3. SVMを用いた滞在地点自動判別

本研究でおこなったSVMによるGPSデータの自動判別について，SVMの概要説明，SVMで利用するテストデータの作成方法，およびSVMを用いた自動判別の結果の説明をおこなう．

(1) Support Vector Machine(SVM)の概要

Support Vector Machine (SVM)は，1995年にAT&TのV.Vapnikが提案したパターン識別用の教師あり機械学習方法である[7]．マージン最大化というアイデアなどで汎化能力も高く，優秀なパターン識別能力を持つとされ，広く研究などに利用されている．SVMは，基本的には与えられたデータを2つのグループに分類する手法である．

SVMの学習データは，学習するデータが属するクラスとデータの特徴ベクトルによって表される．学習データから得られた結果から，属しているクラスがわからない別の特徴ベクトルのデータを入力された時に，どのクラスに属しているのかを判別する．本論文では，GPSの座標データと観測時間の情報などから，入力されたGPSデータが滞在地点であるのか，移動している時に観測されたデータであるのかを判別することになる．

本研究で，用いたSVMの式は次の式(1)で表される．

$$\min_{w}\frac{1}{2}w^{T}w+\sum_{i=1}^{l}(\max(0,1-y_{i}w^{T}x_{i}))^{2} \qquad (1)$$

lは学習データの個数，y_iとx_iは正解ラベルと特徴ベクトルの組み合わせである．式(1)においてマージンを最大にするwを求める．wのパラメータを利用し，特徴ベクトルから2つのグループへの判別をおこなう．

(2) SVMで利用するデータ

本論文で利用する学習データについて，説明をおこなう．本研究では，函館交通行動調査で収集されたGPSデータについて，ある地点に滞在している時のデータであるのか，移動している時のデータであるのかを判別するために，時間的に連続するGPSの座標データと時刻データから，判別をおこなうGPSデータの直前の5つと直後の5つのデータの関係から特徴ベクトルを作成した．特徴ベクトルとしては，判別するGPSデータと各点間の距離と速度，また，判別するGPSデータの直前の5つの平均速度と直後の5つの平均速度の変化率を加えた21の要素を用いて判別をおこなう．

次に，SVMの学習に利用するための正解データの作成について説明をおこなう．本研究では，正解データの作成をGoogleEarthを用いておこなった．函館交通行動調査で得られた3人の移動データを正解データの作成に利用した．被験者の各日のGPSデータをGoogleEarth上にプロットし，地図上のある建物にプロットが集中していた場合，集中している各プロットを滞在の正解データとして設定した．また，移動のデータに関しては，移動速度が早く，プロットが集中している2点間を結ぶような連続したプロットを正解データとした．正解データの数は，滞在と移動の両方共およそ600個作成した．図2は滞在データと移動データの例であり，黄色の枠で囲われている赤いプロットを滞在データとし，青い枠で囲まれたプロットを移動データとした．

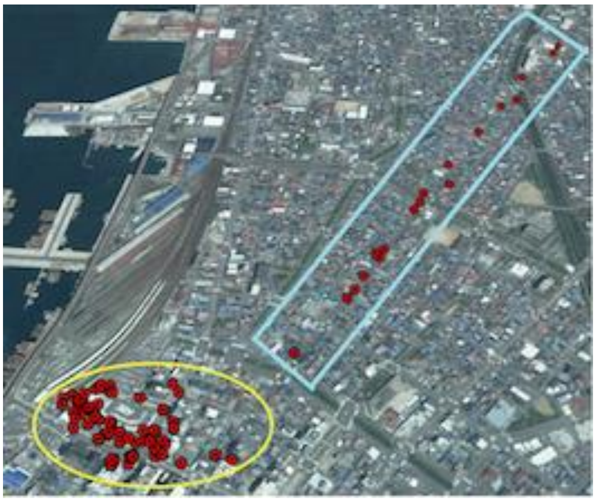

図2 滞在データと移動データの例

(3) SVMによる自動判別

本研究では，学習データから作成したモデルの判別精度を交差検定(Cross Validation)によって確認した．交差検定とは，学習データを複数のグループに分割し，あるグループを学習には利用せず，残りのグループで判別モデルを生成し，学習に利用していないグループのデータを生成したモデルによって判別し，精度を確認する手法である．例えば，学習データをA，B，Cの3つのグループに分割したとすると，BとCを学習に利用し，Aのデータを判別する．次にAとCを学習にしてBを判別し，最後にAとBを学習に利用する．最終的には，全体の判別結果から精度を計算する．本実験では，例と同じく3つのグループに分割して交差検定をおこなった．交差検定を行った結果としては，正解率98.8%となり，本実験で用いたデータでは判別の精度が高いことが示された．

表3 パラメータ推定結果

	定数項1	定数項2	乗車時間	費用	移動距離
パラメータ	1.21636	-0.00251	-0.05707	-0.13304	-0.76759
t値	3.75	-0.00	-4.05	-0.30	-1.56

4. 交通手段選択モデル作成

本章では，交通手段選択モデル構築のために収集したサービスレベルのデータ，およびパラメータ推定の結果について示す．

(1) 交通手段選択モデル

交通手段選択モデルでは，出発地点から目的地点まで移動をする際に，徒歩，自動車，バス，および電車のそれぞれの移動手段にかかる費用や時間などから，どの移動手段を選択するのかを決定するモデルである．スマートシティはこだてプロジェクトでフルデマンドバスを運行した際のシミュレーションで利用されるモデルとなる．

(2) サービスレベルの設定

交通手段選択モデル作成のために，函館交通行動調査で得られたデータに加えて，実際には選択されなかった移動手段の情報を収集する必要がある．交通手段選択モデルの作成では，徒歩やバスなど複数の選択肢がある中で，選択された移動手段と他の移動手段の情報の違いから選択した理由を推定する必要がある．函館交通行動調査では，出発地点から目的地点までどのくらいの時間でどの移動手段で移動したのかという情報は収集されているが，選択されなかった移動手段の情報や，公共交通の費用の情報などは存在しない．そこで，本研究では，実際の移動と同じ出発地点，目的地点，および出発時刻の各サービスレベルの移動の情報を収集した．

サービスレベルの情報の収集にはGoogleMapsとMapionを利用した．GoogleMapsでは，徒歩と自動車の移動にかかる時間と距離を収集した．また，自動車の移動に関しては，15km/リットルの燃費の車を想定し，ガソリン代を150円/リットルと設定して，移動距離からかかる費用を計算した．Mapionでは，公共交通による移動にかかる費用，距離，および乗車している時以外の時間を収集した．乗車している時以外の時間は，待ち時間やバス停までの移動時間などである．徒歩と自動車の移動に関しては，待ち時間は存在しないものとして，0に設定されている．フルデマンドバスの運行においては，通常の公共交通と違い，バス停までの移動時間や待ち時間がほとんどなくなる．したがって，本研究では，待ち時間を重要な情報として設定し，各移動手段の要素として追加した．函館は天候によってサービスレベルが急転する可能性があるが（雪による影響など），本研究ではこれらの要因は加味できていない．サービスレベルのより詳細な考慮は今後の課題である．

(3) パラメータ推定

移動手段選択モデル構築のためのパラメータ推定について述べる．パラメータ推定に用いたデータは，函館交通行動調査で得られたデータを利用しており，サンプル数は380のデータである．移動手段選択モデルの作成では，実際の移動で利用した移動手段の情報が必要になるが，本研究では，移動手段はスマートフォンアプリケーションを利用して収集したアンケートの結果から，実際に利用した移動手段の情報を得ている．

表3はパラメータ推定の結果である．パラメータ推定の結果，$\rho 2$値は0.51となり，モデルの説明力は高い．また，乗車時間は5%有意であるが，費用や移動距離は説明力が弱く，今後，モデルの精緻化が必要である．

5. まとめと今後の課題

本論文では，スマートシティはこだてプロジェクトでシミュレートするフルデマンドバス運行のシミュレーションのためのモデル構築に向けたデータ解析について述べた．2012年12月からおこなわれた函館交通行動調査調査の説明をし，実際に収集されたデータについて示した．

シミュレーションのモデル構築のために必要なデータとして，函館交通行動調査で収集されるデータに加えて，滞在地点のデータが必要なことを示した．Support Vector Machineを用いた滞在地点の自動判別を提案し，実験により高い判別精度を確認した．

スマートシティはこだてのシミュレーションで利用するモデル構築について，函館交通行動調査で収集されたデータとGoogleMapsとMapionから得られたデータを利用して移動手段選択モデルを作成した．

今後の課題として，SVMを利用した利用交通手段の自動判別がある．携帯端末の電池消費量に配慮すると，滞在地点の判別と同様にGPSデータのみから算出できる移動速度などから移動手段を自動判別する必要があると考えている．

被験者が入力するアンケートの結果について，函館交通行動調査では一日の移動を始める前に移動の予定を入力する形式になっていた．そこで事前に入力した移動の予定と実際の移動の違いから，どのような場合に人が移

動手段などの予定を変更するのかといった分析をおこなう予定である.

函館交通行動調査の特徴の一つとして，従来のパーソントリップ調査ではおこなっていない，長期に渡る行動調査があげられる．そこで，本論文の移動手段選択モデル構築では考慮することができていない，人の行動の長期パターンについても分析をおこなっていく必要があると考えている．具体的には，季節，天候，および曜日などの違いによって，人の行動にどのような影響があるのかを調べていく予定である.

謝辞：本研究の一部は，科学技術振興機構社会技術研究開発センター(JST-RISTEX)の問題解決型サービス科学研究開発プログラム’ITが可能にする新しい社会サービスのデザイン”の研究助成によるものである．ここに記して謝意を表する．また，本研究におけるパーソントリップ調査を運用していただいている特定非営利活動法人スマートシティはこだての関係者の皆様，およびパーソントリップ調査に協力していただいている方々に感謝致します.

参考文献

1) 中島 秀之, 車谷 浩一, 伊藤 日出男: ユビキタス情 報処理による社会支援, 情報処理学会誌, Vol. 45, No. 9, pp. 907–911 (2004)
2) 野田五十樹,篠田孝祐,太田正幸,中島秀之: シ ミュレーションによるデマンドバス利便性の評価, 情報処理学会論文誌, Vol. 49, No. 1, pp. 242–252 (2008)
3) 金森 亮,森川 高行,山本 俊行,三輪 富生: 総合交通 戦略の策定に向けた統合型交通需要予測モデルの 開発, 土木学会論文集 D, Vol. 65, No. 4, pp. 503– 518 (2009)
4) 中島 秀之, 白石 陽, 松原仁: 「スマートシティはこ だて」の中核としてのスマートアクセスビークルシ ステムのデザインと実装, 観光情報学会誌, Vol. 7, No. 1, pp. 1–9 (2011)
5) 国土交通省:第 2 回函館圏都市交通調査・都市計画調査 http://www.mlit.go.jp/crd/tosiko/pt/city/hakodate/02.html(2013.3.1)
6) 国土交通省: PT 調査とは http://www.mlit.go.jp/crd/tosiko/pt.html (2013.3.1)
7) C・Cortes，V・Vapnik: Support vector networks，Machine Learning，Vol. 20，pp・273－297 (1995)

(2013. 8. 2 受付)