

卒業論文 2004年度(平成16年度)

MUGI:利用者の行動履歴と場所の属性を用いた
提案型ナビゲーションのための行動モデル

指導教員

慶應義塾大学環境情報学部

徳田 英幸

村井 純

楠本 博之

中村 修

南 政樹

慶應義塾大学 環境情報学部

松倉 友樹

matsu@ht.sfc.keio.ac.jp

MUGI:利用者の行動履歴と場所の属性を用いた 提案型ナビゲーションのための行動モデル

論文要旨

本論文では、行動の習慣性を表現する MUGI モデルを提案し、MUGI を構築する CHOCO システムについて述べる。本研究では、人が行動する際に滞在する場所の位置情報と属性を利用して行動モデルを構築する。場所の属性を利用した行動モデルは、訪れたことのない地域におけるユーザの行動予測を可能にする。

近年、ユビキタスコンピューティング環境の発達により、GPS や RF タグなどのセンサを用いてユーザの位置や行動を取得できるようになった。これにより、システムはユーザの位置や行動を分析し、ユーザの意図や習慣を抽出した行動モデルを構築できる。システムが行動モデルを利用するとユーザの状態に適応したサービスを提供できる。例えば、ユーザが訪問しようとする場所へのナビゲーション情報や、訪問しようとしている場所の情報をユーザの指示無しに提示できる。

しかし、従来の行動モデル構築手法では、行動履歴のみを用いて行動モデルを構築するため、ユーザは行動履歴が存在しない地域では行動モデルを利用したサービスを楽しむことができない。本研究では、ユーザは既知の地域、訪れたことのない地域に関わらず行動に習慣性があると仮定し、訪れたことのない地域で利用可能な行動モデルを構築する。

本研究では、位置情報だけでなく場所の属性を用いた行動モデル MUGI を提案し、MUGI を構築する CHOCO システムを構築する。MUGI はユーザが滞在した場所の属性と、場所から場所への遷移を隠れマルコフモデルを用いてグラフ化する。CHOCO はユーザの習慣性を利用した行動予測と、ナビゲーション経路の評価機能を提供する。アプリケーションは CHOCO の利用により、訪れたことのない地域で行動の習慣性に合ったサービスを提供できる。

キーワード:

行動モデル, 行動履歴, 隠れマルコフモデル, HMM, 位置情報, GIS, GPS

慶應義塾大学 環境情報学部
松倉 友樹

概要

MUGI: Design and Implementation of Activity Models for Proposal Navigation Services using Location Information and Places' Attributes

Summary

This thesis proposes an Activity Model, MUGI, Which represents the habitual aspect of activities, and discusses the construction of CHOCO, a system used in the implementation of MUGI. CHOCO attains attributes from locations we reside in during daily activities, and regards transitions in physical locations as transitions of attributes in the locations, and represents those transitions via the Hidden Markov Model. As a result, users are able to receive services that utilize the Activity Model even in never visited regions.

In recent years, the progression of the ubiquitous computing environment has allowed us to acquire user locations and activities using such sensor devices as GPS and RF tags. Systems can analyze the users' locations and activities, and eventually assemble Activity Models that extract user habits and intentions. With the employment of this Activity Model, such actions as displaying navigational data of user destinations and destination information can be carried out without any user intervention. However, given the traditional Activity Model construction method only takes into consideration user activity records, users are unable to use services based on Activity Models when placed in never visited regions. Yet users, regardless of the level of familiarity with the locations, can be assumed to base his/her actions upon habits. Therefore, it can logically be derived that Activity Models built in known locations can also be applied to unknown locations.

This thesis proposes an Activity Model, MUGI, and implements CHOCO to construct MUGI, an Activity Model which takes into consideration both location information and location attributes, in order to resolve the problems faced by Activity Models dependent upon information from known locations. CHOCO takes the users' location attributes and location transitions data and apply them to a graph using the Hidden Markov Model. It also offers user activity prediction function, using users' habitual data. With the utilization of CHOCO, applications can present services in never visited regions suitable to user habits.

Keywords:

Activity model, Hidden Markov Model, HMM, Location Information, GIS, GPS

Keio University Faculty of Environmental Information
Yuki Matsukura

目次

第1章	序論	1
1.1	背景	2
1.1.1	位置取得センサの普及	2
1.1.2	行動モデルを利用したサービス	2
1.1.3	既存の行動モデル	3
1.2	問題意識	4
1.3	本研究の目的	4
1.4	本論文の構成	4
第2章	既存の行動モデル	6
2.1	行動モデル構築の流れ	7
2.2	データ収集	7
2.2.1	行動履歴	8
2.2.2	動作履歴	10
2.2.3	アンケート	11
2.3	データ分析	11
2.3.1	コンテキスト抽出	11
2.3.2	データマイニング	12
2.4	行動モデル構築	12
2.4.1	確率モデル	13
2.4.2	エージェントモデル	13
2.5	本章のまとめ	14
第3章	場所の属性を用いる行動モデル	16
3.1	目的とアプローチ	17
3.2	データ収集	18
3.2.1	位置情報の粒度と精度	18
3.2.2	行動状態の取得	18
3.3	データ分析	19
3.3.1	滞在場所の特定	19
3.3.2	場所の属性定義	19
3.4	行動モデル構築: MUGI	20
3.4.1	行動遷移の表現	21
3.4.2	行動の学習	22

3.5	本章のまとめ	23
第4章	CHOCO の設計	24
4.1	設計概要	25
4.1.1	ハードウェア構成	25
4.1.2	ソフトウェア構成	26
4.1.3	ソフトウェア動作順序	26
4.2	データ収集部	27
4.2.1	位置取得部	27
4.2.2	位置履歴保存部	28
4.3	データ分析部	28
4.3.1	行動状態判定部	28
4.3.2	場所の属性取得部	28
4.3.3	場所の属性保存部	28
4.4	行動モデル構築部	28
4.4.1	行動モデル管理部	29
4.4.2	行動モデル部：MUGI	29
4.5	アプリケーションインタフェース	29
4.6	本章のまとめ	30
第5章	CHOCO の実装	31
5.1	実装環境	32
5.1.1	コンピュータ環境	32
5.1.2	位置取得システム	33
5.1.3	ソフトウェア設計	33
5.2	データ収集部	33
5.2.1	位置取得部	34
5.2.2	位置履歴保存部	34
5.3	データ分析部	35
5.3.1	行動状態判定部	35
5.3.2	場所の属性取得部	35
5.3.3	場所の属性保存部	36
5.4	行動モデル構築部	37
5.4.1	行動モデル管理部	37
5.4.2	行動モデル部：MUGI	37
5.5	アプリケーションインタフェース	38
5.6	サンプルアプリケーション	39
5.7	本章のまとめ	40

第 6 章	CHOCO の評価	41
6.1	定量的評価	42
6.1.1	学習の収束度	42
6.1.2	スケーラビリティ	44
6.2	定性的評価	45
6.3	本章のまとめ	45
第 7 章	結論	46
7.1	今後の課題	47
7.1.1	習慣性の変化に対応	47
7.1.2	フィードバック機構	47
7.1.3	行動モデルの共有	47
7.2	まとめ	47
付 録 A	行動状態の判定	51

目次

1.1	行動モデルを利用したロケーションウェアサービス	3
2.1	行動モデル構築の手順	7
2.2	ハンディタイプの GPS 受信機	8
2.3	位置情報の用語定義	9
2.4	画像解析を用いた位置取得手法	10
2.5	動作履歴と位置情報の関連を可視化	10
2.6	行動履歴を地図上に表示	12
2.7	ベイジアンネットを用いた行動分析	13
2.8	滞在点を示した行動モデル	14
2.9	マルチエージェントモデルを用いた行動シミュレーション	15
3.1	既存研究 (左) と本研究 (右) のデータの流れ	17
3.2	隠れマルコフモデルを用いたモデル	22
4.1	ハードウェア構成	25
4.2	ソフトウェア構成	26
4.3	アクティビティ図による CHOCO の動作順序	27
5.1	ハードウェア環境	32
5.2	位置取得部クラス図	34
5.3	位置履歴保存部クラス図	35
5.4	行動状態判定部クラス図	36
5.5	PostGIS を利用した SQL による問い合わせ例	36
5.6	場所の属性取得部クラス図	37
5.7	行動モデル保存部クラス図	38
5.8	CHOCO の利用例	39
5.9	HMM をグラフィカル表示	39
5.10	HMM をの内部属性を表示	40
6.1	使用した行動履歴を地図にプロット	42
6.2	滞在数に対するノード数の変化	43
6.3	滞在数に対するリンク数の変化	43
6.4	滞在数に対するパラメータ数の変化	44
6.5	滞在回数に対するファイルサイズの変化	45
A.1	平均移動距離	52

表 目 次

2.1	位置取得方法	9
3.1	NMEA 0183 フォーマット抜粋	19
3.2	隠れマルコフモデルの属性との対応	22
4.1	行動評価機能の入出力	29
5.1	実装環境	32
5.2	GPS 受信機仕様	33
5.3	行動モデル三手順とパッケージの対応	33
6.1	用いた行動履歴の詳細	42
6.2	用いた行動履歴の詳細	45
A.1	実際の移動	51

第1章 序論

1.1 背景

近年、情報処理能力および、ネットワーク接続性を持った様々な機器が環境に遍在したユビキタスコンピューティング環境 [18] が実現されつつある。その環境では、ユーザは場所や時間を気にせずにネットワークへ接続でき、いつでもどこでもネットワークを介して情報を送受信できる。そして、環境に設置されたセンサが協調動作し、ユーザ自身またはサービス提供者はユーザの位置情報を得られる。

ユビキタスコンピューティング環境においてサービス提供者は、様々なセンサから取得した情報を利用し、ユーザの位置や行動を取得し、分析することによりユーザのコンテキストを取得できる。コンテキストとは、ユーザが滞在している位置や行動している状況などの背景となる情報である。一般に、計算機器がユーザのコンテキストを利用し、ユーザの状態に応じて提供するサービスはコンテキストウェアサービスと呼ばれる [6]。特に、ユーザの位置に応じて提供するコンテキストウェアサービスはロケーションウェアサービスと呼ばれる。

次に、ユビキタスコンピューティング環境の普及に伴う位置取得センサの現状と、行動モデルの背景を述べる。

1.1.1 位置取得センサの普及

近年、位置取得技術の発展はめざましく、屋内外で位置情報を取得できる環境が整いつつあり、ロケーションウェアサービスを提供できる基盤が整いつつある。

屋外では、GPS (Global Positioning System) を用いることにより、ユーザは現在地を地球上の絶対座標で取得できる。2001年の米国によるGPSの精度制限の解除により、数メートルの精度で位置情報を取得できる [5]。そのため、人がどの建物の前にいるのかを判別できる。また、小型GPSレシーバの登場により、GPSレシーバを搭載した携帯電話 [2] が登場し、ユーザの位置を取得できる環境が整いつつある。

屋内では、能動的に電波を発信する小型の機器であるActive型RFタグや、超音波発信器を用いることによりサービス提供者は目的に応じた粒度と設備で位置を取得できる。Active型RFタグを用いた位置取得方法は、一意の識別子を発信する電池付タグをユーザが持ち、その識別子を部屋や廊下などに取り付けられたリーダーで読み取る。次に、超音波発信器を用いた位置取得方法では、天井や壁などに超音波受信機を設置し、ユーザが超音波発信器を所持することにより、数センチメートルの粒度で位置を取得できる [16]。

これらの位置取得技術はイベント会場のような限られた場所で利用されているのが現状であるが、今後は屋内の至る所でユーザの位置を取得できる環境が整うと考える。

1.1.2 行動モデルを利用したサービス

人の日常の行動には、習慣性が存在すると考える。習慣性とは、位置や時刻などの特定の条件下でその人が起こす行動に規則性が存在することである。例えば、ある人が“レス

トランへ行った後(条件)にカフェへ行く(行動)”という習慣性である。

行動の習慣性を利用すると、日常の生活をサポートする有用なサービスを提供できる(図1.1)。例えば、あるユーザがレストランAの後にカフェBへ行く習慣があり、ユーザがレストランAに居るときに、ユーザが所持しているPDAが行動モデルを用いて次にユーザが訪れる場所を予測し、その場所への最適な経路や、その場所の混雑度や定休日などの情報を提示できる。

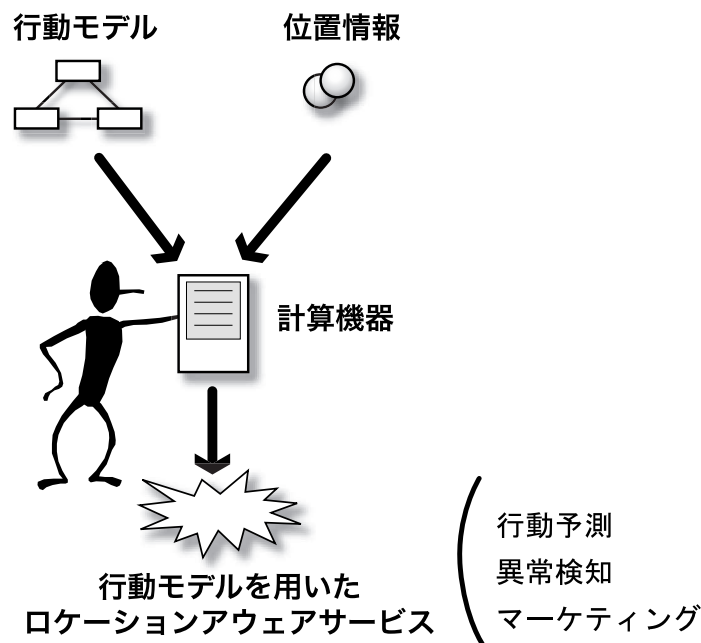


図 1.1: 行動モデルを利用したロケーションウェアサービス

このようなサービスを実現するためには、システムが人の行動履歴を解析し、行動の習慣性を抽出し、ユーザの行動の特徴を表現した行動モデルを構築する必要がある。

1.1.3 既存の行動モデル

過去に、人の行動モデルを構築するための様々な手法が研究されている。既存の研究では人が行動する位置の遷移に注目して行動モデルを構築している。

Inferring High Level Behavior from Low Level Sensors [14] では GPS で取得した行動履歴と地理位置情報を利用し、一人のユーザに対して複数の行動モデルを構築し、状況に応じて行動モデルを切り替える手法を提案している。移動手段ごとに行動モデルを構築することでユーザの行動形態の変化する場所が限定される。例えば、バスの行動モデルでは停留所、電車の行動モデルでは駅が行動の変化する場所である。移動手段によってユーザが行動できる範囲が制限されるので、移動手段に特化した行動モデルを構築している。

A Hybrid Location Model with a Computable Location Identifier for Ubiquitous Computing [12] では GPS で取得した行動履歴を用いて、都市間の移動や会社と家といった地

域間での移動を行動をモデル化する手法を提案している．行動モデルの構築のためにユーザが移動する際の始点と終点を組にして履歴をとり，始点と終点の分布を地理上の地域ごとにクラスタリング分析し，特徴のある遷移行動を抽出している．次に，分類した地域間の移動を確率モデルを用いて行動モデルを構築している．

HMMによる行動パターンの認識 [24] では室内における人の位置をカメラを用いて取得し，人が普段滞在する位置を抽出する．滞在する位置とその位置同士の関連を隠れマルコフモデル (HMM) を用いて表し，室内での行動モデルを構築している．

1.2 問題意識

既存の行動モデル構築手法では，ユーザの行動履歴がない地域ではそのユーザの行動モデルが存在しないため，行動モデルを利用したロケーションウェアサービスを楽しむことができない．既存の研究では行動モデルを構築するために，ある限られた地域において行動履歴を取得し，人が行動する位置の遷移に着目して行動モデルを構築している．従って，ユーザが初めて訪れる地域では，その地域に特化した行動モデルを構築するために新たな行動履歴が必要になる．

よって，ユーザが初めて訪れる地域に居るときに，行動モデルを用いたロケーションウェアサービスを楽しむことができない．

1.3 本研究の目的

ユーザが初めて訪れた地域に居る際も利用できる行動モデル構築手法を提案する．これにより，ユーザは初めて訪れた地域に居る際も行動モデルを用いたロケーションウェアサービスを楽しむことができる．

人の行動は，既知の地域と初めて訪れる地域でも共通の条件があり，その条件に基づいた共通の行動をすると考える．よって，既知の地域で構築した行動モデルを初めて訪れる地域でも応用できると考える．例えば，既知の地域でレストランの次にはカフェに行くという習慣があるユーザに対し，初めて訪れる地域でもユーザの習慣的な行動を支援できる．

本研究では，初めて訪れる地域で利用可能な行動モデルである MUGI を提案し，MUGI を実装したミドルウェアである CHOCO を構築する．ロケーションウェアサービス提供者は CHOCO を利用することで，ユーザの行動モデルを利用したサービスを構築できる．

1.4 本論文の構成

本論文では，第 2 章において，行動モデル構築の手順を分類し，各手順中の要素技術を説明する．第 3 章ではまず，本研究の機能要件を整理し，次に要件を満たす行動モデルを構築するために第 2 章で分類した手順に則り詳しく述べる．第 4 章で初めて訪れる地域

で利用可能な行動モデルである CHOCO の設計について，第 5 章で実装について述べる．
そして，第 6 章で CHOCO を評価し，第 7 章で本論文のまとめを述べる．

第2章 既存の行動モデル

本章では，行動モデル構築に必要な手順を三段階に分類し，各段階における要求を整理する．次に，各段階における既存の手法について述べ，利点と欠点を整理する．

2.1 行動モデル構築の流れ

一般に、人の行動には規則性や習慣性が存在すると考えられる。例えば、コンビニエンスストアでは、お客の行動パターンを分析し、行動パターンに沿う様な商品陳列を行い、購買を促進している。例えば、米国のスーパーマーケットチェーンのウォール・マートでは「週末に紙おむつを買いにくる男性は、同時に缶ビールをケースごと買う」という行動パターンを導き出し、ビールと紙おむつを同時に陳列して売り上げを伸ばした。

ユビキタスコンピューティング環境においては、屋内外にてユーザの位置情報を取得できる。よって、屋内外にとらわれずに行動履歴を収集できるため、屋内外におけるユーザ行動の規則性を取得できる。

私は行動モデルの構築を構築するためには三つの要素と順序がある事を発見した。なぜなら、関連研究の行動モデル構築の順序を分析すると、“データ収集”、“データ分析”、“行動モデル構築”の三要素に分類できる。関連研究を挙げて例示すると、“Learning Significant Locations and Predicting User Movement with GPS[1]”ではデータ収集としてGPSによる行動履歴を取得し、データ分析として行動履歴から特徴点を抽出し、行動モデル構築では特徴点の関連性を確率モデルにより表現している。また、“積雪寒冷地における震災対応マルチエージェントシステムの開発[23]”ではデータ収集としてアンケートによる調査を行い、データ分析としてアンケートからモデルで表現する特徴を抽出し、行動モデル構築では個人の特徴を反映した仮想エージェントを用いて個人行動を表現している。

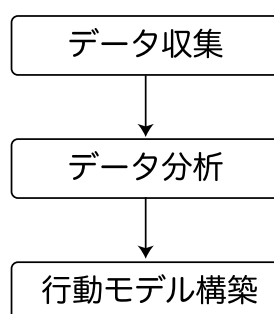


図 2.1: 行動モデル構築の手順

本論文では図 2.1に示す順序で行動モデルを構築する手順を“行動モデル三手順”と定義する。行動モデルの構築には三手順をデータ収集、データ分析、行動モデル構築の順に処理する。

次に関連研究では行動モデル三手順の用いられている手法を例に説明する。

2.2 データ収集

データ収集部は行動モデル構築において、抽象化する特徴を取得するための基礎となる情報を取得する。

データ収集の手法は主に行動履歴，動作履歴，アンケートの三種類がある．次に，各手法についての説明と関連研究を述べる．

2.2.1 行動履歴

行動履歴とは，屋内または屋外で取得したユーザの位置を，時間とともに記録した情報である．

屋外で人の位置を取得するために，地球上の位置を絶対座標で取得できる GPS (Global Positioning System) が利用されている．GPS とは，地球上にある受信機が地球を周回する 24 個の衛星から放射される電波を利用して三点測量を行い，受信機の位置を取得するシステムである．精度は，見通しのよい場所で数メートルの誤差，周囲にビルなどが建ち，見通しの悪い場所では十数メートルの誤差がある．

GPS を利用するためには，GPS 衛星からの電波を受信するための GPS 受信機が必要である．GPS 受信機は長さ 10cm 程度のハンディタイプ (図 2.2) がガーミン社 [9] などから市販されている．

GPS は “Learning Significant Locations and Predicting User Movement with GPS[1]” や “Inferring High Level Behavior from Low Level Sensors[14]” にて，広い範囲にわたるユーザの行動履歴を取得するために利用されている．



図 2.2: ハンディタイプの GPS 受信機

屋内で人の位置を取得する手法は多種多様であり，超音波センサ，無線 LAN，動画カメラ，無線タグが利用されている．超音波や無線 LAN を用いた位置取得手法では，建物側に基地局を設置し，人の持つ受信機が基地局から発せられた音波や電波強度を用いて三点測量して位置を取得する．超音波を用いた測量では数センチメートルの精度で位置を取得できる．RF タグは “滞在時間を考慮した主要行動パターン抽出方法の検討 [19]” にて，展示会の来場者の行動を取得するために利用されている．画像解析を用いた位置取得手法では，建物側にカメラを設置する．そのカメラで定期的に写真を撮り，回数 c における画像 $P(c)$ と $P(c-1)$ とのフレーム間差分をユーザの移動とする (図 2.4) ．画像解析は “Learning and Recognizing Behavioral Patterns Using Position and Posture of Human[11]” で利用されている．

屋外での位置取得技術は GPS があり，屋内では RF タグや超音波を使った位置取得技術がある．それぞれの粒度・導入コストを表 2.1 に示す．

表 2.1: 位置取得方法

適応場所	粒度	導入コスト	位置取得技術
屋外	細かい	低い	GPS
屋内	荒い	低い	RF タグ・無線 LAN
	細かい	高い	超音波・画像解析

位置情報を表記するために位置や範囲を表記する用語は様々な種類があり，その用語の理解には誤差がある．例えば，位置情報を表す用語は，位置・ロケーション・場所・地帯・ポリゴン・地域などであり，場所や地域が表す範囲は用いる際の文脈によって変化する．本論文で用いる位置情報を表す用語とその定義を次に示す．

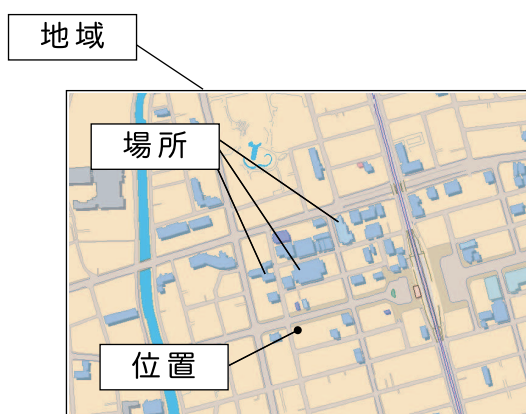


図 2.3: 位置情報の用語定義

- 位置

位置は，2次元座標において点を表す．位置は GPS など取得する座標である．例えば，“北緯 35 度 23 分 7 秒 東経 139 度 25 分 27 秒” は位置である．

- 場所

場所は，2次元座標において多角形の線で囲める範囲を表す．例えば，“北緯 35 度 23 分 7 秒 東経 139 度 25 分 27 秒 から 10 メートル四方の範囲” は場所を示す．実世界では，商店，ビル，公園，学校などの建物程度の範囲を示す．

- 地域

地域とは，2次元座標において多角形の線で囲める範囲を表し，地域は複数の場所を包括する．例えば，実世界では，市町村，街区，駅前などの広範囲を示す．

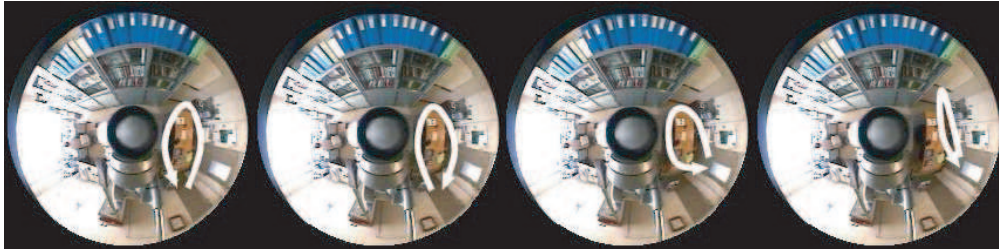


図 2.4: 画像解析を用いた位置取得手法

2.2.2 動作履歴

動作履歴とは、環境に設置された機器に対してユーザが起こした対象機器と動作を起こした時間の記録である。また、動作履歴を保存するためには、ユーザが操作する機器は計算機能を備えてなければならない。例えば、電子マネー機能を持った JR 東日本の“Suica[®]” や関東の私鉄で利用できる“パスネット[®]” の場合、駅の改札で入退場記録とユーザを関連づければそのユーザの電車の乗り降りの動作履歴を保存できる。

動作履歴は“A Behavior-based Personal Controller for Autonomous Ubiquitous Computing[13]”にて、行動予測のために用いられている。その論文では、ユーザの動作履歴と同時に位置情報を記録し、図 2.5 に挙げるグラフに表す。図 2.5 にある点は動作を表し、点のある位置はユーザの位置を表してユーザの行動モデルを表現している。行動予測の手法は、ユーザの現在地をグラフ上にプロットし、その位置から一番近い動作を求める。

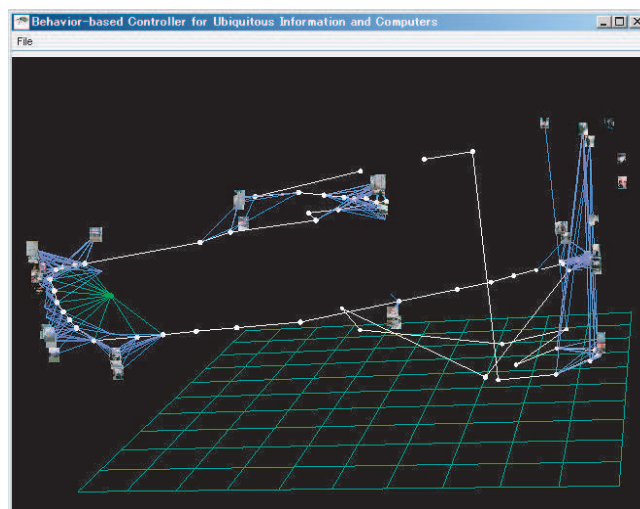


図 2.5: 動作履歴と位置情報の関連を可視化

2.2.3 アンケート

ある決められた指標に基づいた質問をユーザが回答し、ユーザの情報を取得する手法。アンケートの欠点の一つは、アンケートの文章や回答項目によって得られる回答が異なる。そのため、アンケートによる正確なデータ収集が困難である。アンケートによる手法は、多数のユーザによって回答してもらい、その解析結果を利用するのが一般的である。アンケートによるデータ収集は、“ベイジアンネットワークを用いた消費者行動モデルの構築実験 [22]” にてユーザの意識調査をするために用いられている。

2.3 データ分析

データ分析部は、行動モデル構築で利用するデータへ収集したデータを加工する。データ収集部で収集されたデータは行動履歴やアンケート結果といった未加工のデータであるため、行動モデルで利用しやすいフォーマットへ変換、または特徴の抽出を行う必要がある。また、データ収集時にセンサを用いた場合は、正確なデータ分析を行うために、はずれ値の除外やデータの正規化を行う。それにより、行動モデルを構築に利用する情報がはずれ値などを含まない情報に改善される。

データ分析の手法は主にコンテキスト抽出、データマイニングの2種類がある。次に、各手法についての説明と関連研究を述べる。

2.3.1 コンテキスト抽出

コンテキスト抽出とは、実際の行動を元に未加工のデータの中からその行動と相関性のある特徴を見つけることである。

行動履歴を例にとると、収集した行動履歴にはユーザが用いている行動手段や滞在している場所の情報は含まれていない。行動モデルを構築するための情報を行動履歴からの抽出や、行動履歴と他の情報を組み合わせて行動の特徴を抽出する。

“Learning Significant Locations and Predicting User Movement with GPS[1]” ではユーザの行動の特徴点として、GPS で取得した行動履歴から滞在した地点を自動的に抽出している (図 2.6)。

“Inferring High Level Behavior from Low Level Sensors[14]” では、GPS で取得した行動履歴からユーザのあらゆるコンテキストを取得し、行動予測を行うことを目的としている。この研究ではベイジアンネットワーク [8] を用いてユーザの行動状態の判別を行っている。ベイジアンネットワークは確率変数を表すノードと、確率変数間の直接的な依存関係を表す枝により定義される有向無サイクルグラフである。グラフの構造により確率モデルの構造が指定され、各ノードに割り当てられた条件付確率分布 (Conditional Probability Distribution:CPD) によりモデルのパラメータが表現される。図 2.7では、GPS の位置情報から算出した速度と、バス停や駐車場の場所を用いてユーザの行動状態を予測するためのベイジアンネットワークである。



図 2.6: 行動履歴を地図上に表示

2.3.2 データマイニング

データマイニングとは、未加工のデータの中に存在するある傾向や相関関係などの情報を見付け出すための手法である。アンケートで取得した回答を元に、アンケート回答者の性格などを分類する。

“積雪寒冷地における震災対応マルチエージェントシステムの開発 [23]” では、積雪の多い地域の住民にアンケートを行い、住人の行動を指標化している。

2.4 行動モデル構築

行動モデル構築部ではデータ分析によって抽出した行動の特徴やユーザの特徴を確率や人工知能を用いて表現する。行動モデルはユビキタスコンピューティングの分野に限らずマーケティングでも用いられている。

行動モデルの表現手法には確率モデルとエージェントモデルがある。確率モデルを用い

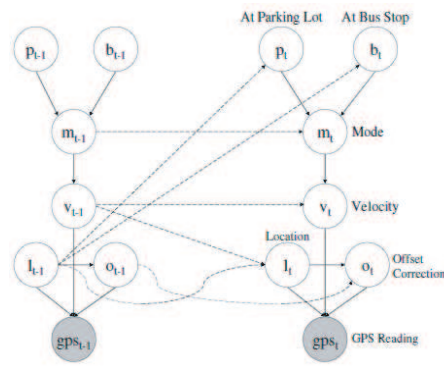


図 2.7: ベイジアンネットを用いた行動分析

る場合は、環境が変化したときのユーザの行動をモデルする場合に用いられ、エージェントモデルは、ある一定の環境下でのユーザ行動を表現する場合に用いられる。

2.4.1 確率モデル

確率モデルとは、一般的に要素と要素をグラフで表現し、確率によって要素との関連を表現したグラフである。

行動モデル構築で確率モデルを適用する場合は、対象となる事象である、不確定な人の動きという要素を確率として扱い、不確定な要素と関連性のある要素との関連性を確率で表現する。

また、確率モデルは、行動モデルとして表現する事項や予測する事項によって様々なアルゴリズムが提案されている。例えば、ベイジアンネットや隠れマルコフモデル [15] がある。

2.4.2 エージェントモデル

エージェントモデルとは、計算機器上に仮想的に構築した環境で自律的に行動するエージェントを動作させるモデルである。マルチエージェントモデルは、エージェントとエージェントが行動する環境により構成される。エージェントにはデータ分析により得られた人の特徴などを属性として持たせ、仮想環境で人がどのような行動を取るのかを観察する。

また、エージェントを膨大な数作成し、環境とエージェントだけでなくエージェントの相互作用もモデル化したマルチエージェントモデルが用いられる。

エージェントモデルの利点は、エージェントは人自体をモデル化しているため、仮想的に構築された環境において人がどのような行動をするのかといった事象を検証できる。欠点は、エージェントと動作環境のモデル化に大きく左右されるため、検証する事象に関連するエージェントと動作環境の表現が難しい。

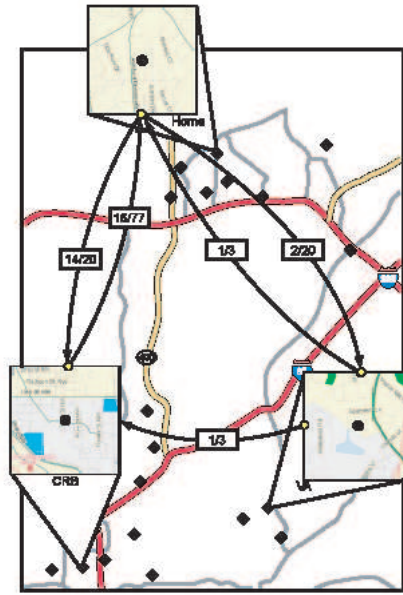


図 2.8: 滞在点を示した行動モデル

“積雪寒冷地における震災対応マルチエージェントシステムの開発 [23]” では、ある地域で災害が起きた際の人間の行動を観察するためにマルチエージェントモデルを用いている。積雪のある地域に住む住民にアンケートをとりユーザの行動特徴をエージェントの属性として持たせ、エージェントが動作する環境とし、その地域の地図と経路を用意する(図 2.9)。シミュレーションでは災害が起きたことを仮定し、ある経路を遮断したときのエージェントの行動などを観察する。

2.5 本章のまとめ

本章では、行動モデル構築のための手順を三手順に分類し、各段階における要件を述べ、各段階における既存のアプローチについて述べ、利点と欠点を整理した。

次章では、訪れたことのない地域で行動モデルを利用するための要件を、本章で定義した行動モデル構築三手順ごとのアプローチを述べる。

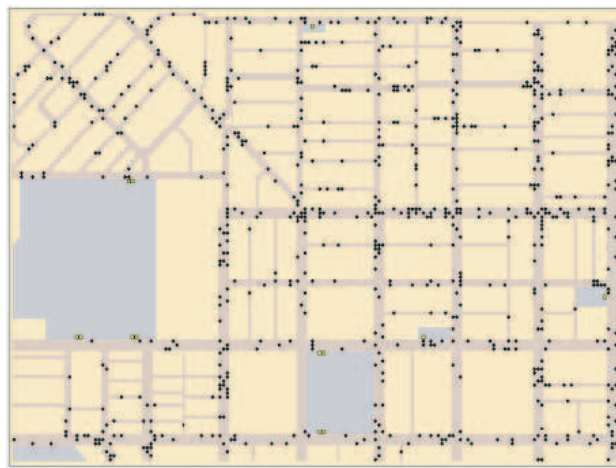


図 2.9: マルチエージェントモデルを用いた行動シミュレーション

第3章 場所の属性を用いる行動モデル

本章では，未知の地域で利用できる行動モデルを構築するための機能要件を整理し，行動モデル構築のアプローチを述べる．

3.1 目的とアプローチ

本研究の目的は、ユーザが未知の地域へ訪れた際も利用できる行動モデル構築手法を提案することである。これにより、ユーザは、未知の地域でも行動モデルを用いたロケーションウェアサービスを楽しむことができる。

本研究のアプローチでは、未知の地域で利用できる行動モデルを構築するために、行動履歴から取得した位置情報のみを用いるのではなく、“場所の属性”も利用して行動モデルを構築する。場所の属性とは、その場所をある規則に従って分類した場所のメタ情報である。行動モデルに利用する場所の属性は、ユーザが滞在した場所の情報をモデル化する。

既存研究では位置の座標値を解析し行動モデルを構築している。そのため、行動モデルを構築するためにはその地域での行動履歴が必要となる(図3.1左)。本研究では、訪れた場所の位置座標に対応した場所の属性を取得し、属性間の遷移をモデル化する(図3.1右)。その結果として、ユーザの位置情報と行動モデルの結合度が低くなり、ある場所で構築した行動モデルを別の場所でも利用できる。

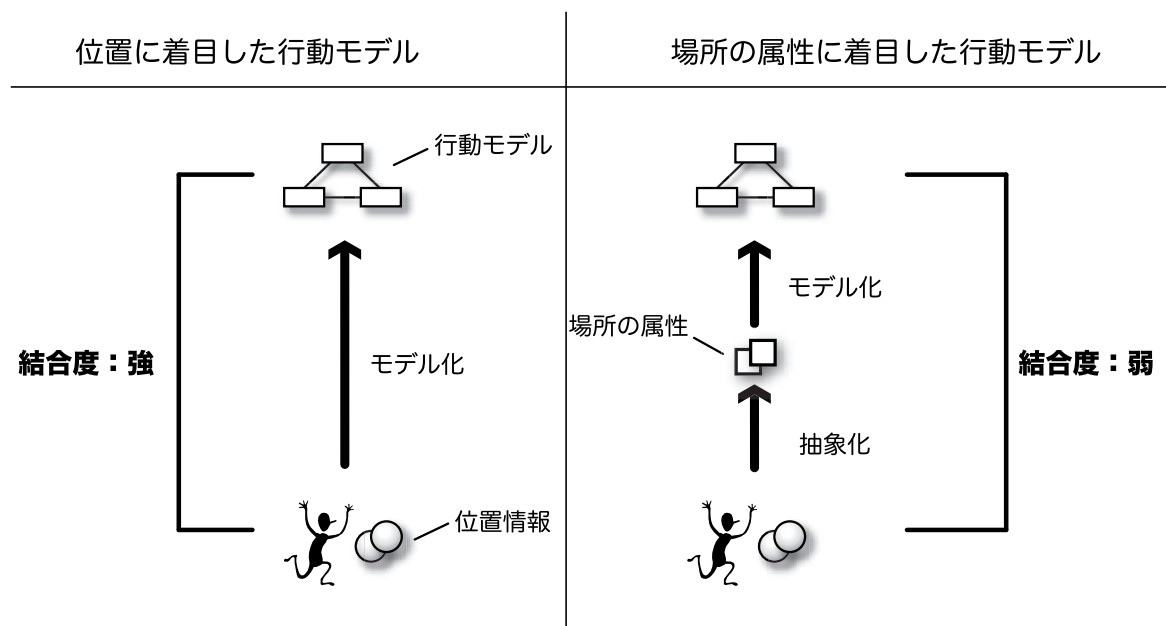


図 3.1: 既存研究 (左) と本研究 (右) のデータの流れ

第2章で述べた行動モデル構築三手順について本研究への適用を念頭にその利点および欠点を整理する。

3.2 データ収集

行動モデル構築のための基礎となるデータを取得する。アンケートを用いてデータを収集した場合には、新たな行動モデルを構築する場合には新たにアンケートを採らなければならないため不適切である。また、動作履歴を用いる手法は現在のコンピューティング環境の状態から、実現性が低いと考える。そのため、本研究では行動履歴を用いてデータを収集する。以下に行動履歴を用いた際の機能要件を述べる。

- 行動履歴の粒度と頻度

位置情報は、場所の属性を取得するために、ユーザの場所を一意に特定できる必要がある。そのための位置取得システムの要件を、以下に示す。

- 粒度

ユーザがどの場所に滞在しているかを取得するために、商店街などの店舗が密集した地域でどの店舗に滞在しているかが判別できる必要がある。具体的には、約5メートル以上の細かさが必要になる。

- 頻度

ユーザの行動常態である停止・動作の状態を取得するために十分な頻度の履歴を取得する間隔が必要である。間隔が長いと一定時間内の履歴数が少なくなり、はずれ値を除外するための条件設定が困難になる。

- 滞在状態の取得

ユーザが滞在した場所の属性を取得するために、その場所に滞在しているかどうかを判別する必要がある。滞在とは、行動の種類は動作か停止の二つに分類できる。滞在とは、ユーザがある場所に一定時間居ることと定義する。

3.2.1 位置情報の粒度と精度

屋外で上記の要件を満たす位置取得方法としてGPSを用いる。GPSは数メートルの精度があるため、ユーザがどの場所にいるかを特定できる。

また、GPSは誤差を出力できるため、天候や遮蔽物の影響で位置取得精度が悪い場合は値として採用しないようにできる。

3.2.2 行動状態の取得

行動状態は、GPSから取得した位置情報から分析する。GPSが出力する測位結果の規格として米国海洋電子機器協会(National Marine Electronics Association)[3]が定義した“NMEA 0183 Standard”[4]がある。NMEA 0183は市販のGPS受信機の多くはこのNMEA 0183の出力に対応している。NMEA 0183で規定されている出力形式の、位置情報に関する部分を抜粋する。

項目	値
観測時刻	位置を取得した時刻
緯度	北緯/南緯 度分
経度	西経/東経 度分
使用衛星数	測位に使用した衛星数
水平測位誤差	緯度経度の誤差
ジオイド高	海水面基準からの高さ
海拔高度	ジオイド高からの垂直高さ

表 3.1: NMEA 0183 フォーマット抜粋

3.3 データ分析

行動モデル構築のために必要な解析を行う。MUGIを構築するためにはユーザの滞在した場所の属性が必要になる。データ分析部ではユーザが滞在したかどうかを判断し、滞在した場所の属性を取得する。取得した位置情報からの滞在している場所を判別、場所の属性の定義の要件を以下に示す。

- 滞在場所の特定

取得した現在の位置情報と位置履歴を参照し、ユーザが滞在しているかを判定する。滞在の状態とは、滞在する目的があつて行った場所である。例えば、食事をするために行ったレストランは、滞在している場所である。それに対し、信号待ちのために停止した状態は滞在ではない。

- 場所の属性定義

滞在している場所の分類は全国で一意に表現される必要がある。地域によって場所の定義方法が異なると、同じ場所が別の属性と認識してしまう。

3.3.1 滞在場所の特定

ユーザが滞在している場所の属性を取得するためにある一定時間ユーザが停止した場合にその場所に滞在したと判定する。その場所に目的があつて滞在したのか、それとも信号待ちのために滞在したのかが不明である。そこで、ユーザの停止位置と停止地点からの建物の距離を参考に滞在場所の属性を取得する。

3.3.2 場所の属性定義

場所の属性はユーザ共通の属性とユーザ独自の属性がある。ユーザ共通の属性とは、レストランやカフェなどの地図に記載されている全てのユーザに共通する情報である。ユー

ザ独自の属性とは，地図に記載されていない属性であり，自宅や通学している学校などのユーザごとに異なる情報がある．次に，それぞれの属性の取得・定義方法について述べる．

- ユーザ共通の属性

場所の属性は様々な定義をできる．例えば，A 駅のカフェを説明する属性は，A 地区，店舗，飲食店，カフェ等の属性が考えられる．

地図上の場所や位置に情報を持たせるために地図を電子化した，地理情報システム (Geographic Information Systems) がある．GIS では空中写真や統計データなどの様々なデータを層に分けて管理し，位置をキーに地図上に位置や場所関連づけるシステムである．GIS の整備は，政府が取り組んで行っており，様々な地理位置情報を Web からダウンロードして利用できる [20]．本研究では，場所の属性を GIS から取得して利用する．

- ユーザ独自の属性

ユーザ独自の属性とは，ユーザが頻繁に行く場所であり，市販されている地図や GIS には記載されない属性である．自宅は地図などでは家屋と表記されるが，個人にとっては自宅という意味のある場所である．ユーザ独自の属性の場所は，自宅・学校・職場がある．

ユーザ独自の属性は自宅や学校などのように，その地域に依存した場所であるため，未知の地域ではこの属性は行動モデルでは利用しない．

3.4 行動モデル構築: MUGI

データ分析部より取得した場所の属性を利用して行動モデルを構築する．本研究では，ユーザが行動した場所の属性を用い，移動した場所の属性間の関連性をモデルとして表現する．

既知の場所での行動において場所の属性を学習し，その関連を未知の地域で適応できる MUGI(A Model for User activity using Geographical Information) を提案する．次に MUGI の機能要件を述べる．

- 行動遷移の表現

ユーザの行動の習慣性を抽出するために，行動履歴を用いる．継続的に増加する行動履歴を扱うためには，長期利用においての保存要領を考慮する．

- 行動の学習

行動モデルの学習時は，ユーザの明示的な入力を求めずにモデルを構築する．モデル構築の際に，ユーザが滞在した場所をその都度入力するのは負担になるため，システムがユーザの位置情報を取得し行動モデルへ反映する．

3.4.1 行動遷移の表現

本節では、属性に注目した行動モデルを表記するためのアプローチを示す。

本研究では、行動モデル構築のために確率的な状態遷移と確率的な出力を備えたオートマトンである“隠れマルコフモデル”(Hidden Markov Model 以降 HMM) を利用する。

HMM を用いる理由は、まず、人が行動する場所の属性に注目し、その移り変わりを表すためには属性と属性の関係を記述する必要がある。そこで一つの属性を“ノード”として表し、ノード間の遷移関係を確率で示す。

また、HMM では EM アルゴリズムによる計算が可能であるため、ユーザがモデルの学習を行った直後からサービス提供者は行動モデルを利用したサービスを提供できる。

状態遷移を表すための確率モデルは HMM の他に、ベイジアンネットワーク [8] がある。ベイジアンネットワークはグラフで表現され、確率変数をノードと置き、ノード間の有向リンクで構成される。

しかし、ベイジアンネットワークのノードは循環してはならないという制限がある。よって、ベイジアンネットワークのノードを属性とにおいてグラフを生成した場合には、一度訪れた場所の属性には訪れないという制約が発生するため非現実的である。

次に、HMM の構成要素である、入力記号列、遷移確率、内部状態、出力確率について述べる。

- 入力記号列

行動モデル構築のために、ユーザが行動した場所の属性遷移を入力する。例えば、“属性 A 属性 B” という遷移である。

- 遷移確率

あるノードから、次に遷移するノードへの確率を示す。ユーザの行動に基づいてノードからノードへ遷移する確率が変化する。例えば、ノード A からノード B へ遷移する確率 $P_A(B|A)$ の計算式を次に挙げる。

$$P_A(B|A) = A \text{ から } B \text{ へ遷移した回数} \div A \text{ から遷移した回数}$$

- 内部状態

訪れた場所の名称と回数を正規化して保存する。ノード A から、ノード B の場所 b1 へ遷移する確率 P_1 の計算式を次に挙げる。

$$P_1(b1|B) = P_A(B|A) \times (b1 \text{ に訪れた回数} \div B \text{ へ訪れた合計回数})$$

- 出力確率

ある属性の場所から、次に訪れる場所の属性または次に訪れる場所へ到達する確率を示す。次にノード A からノード B の場所 b へ遷移する確率 P_2 の計算式を挙げる。

$$P_2 = P_A(B|A) \times P_1(b|B)$$

HMM の構成要素を MUGI では表 3.2 の右列のに挙げるように割り当てる .

属性	本研究での割当
入力記号列	行動した場所の属性履歴
遷移確率	現在地から次に遷移する場所の属性への確率
内部状態	行動した場所の名称の履歴
出力確率	ある場所から , 別の場所へ訪れる確率

表 3.2: 隠れマルコフモデルの属性との対応

表 3.2 の割り当てに従い , HMM を用いて表現した行動モデルの例を図 3.2 に示す . この場合では , ノード A からノード B への遷移確率は 0.82 , ノード B の中でも場所 b1 へ遷移する確率は 0.73 ということを表している .

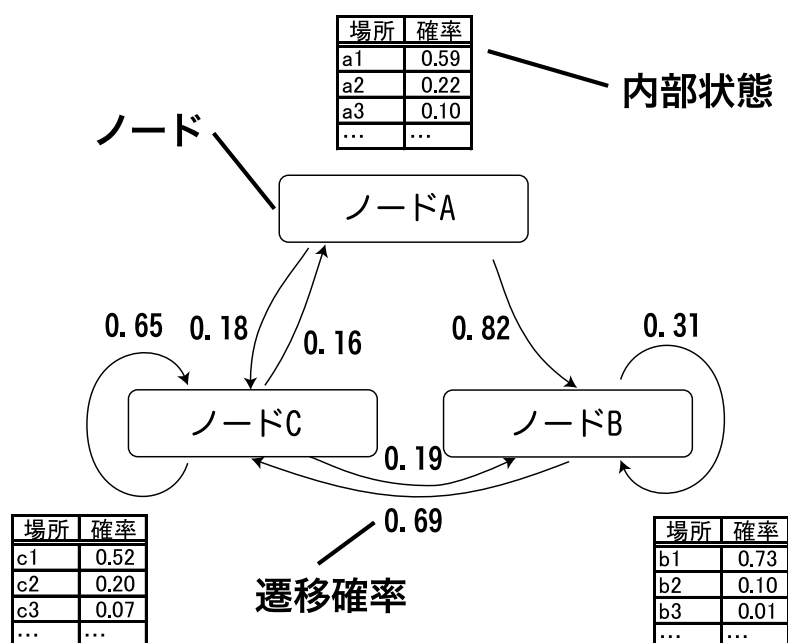


図 3.2: 隠れマルコフモデルを用いたモデル

3.4.2 行動の学習

HMM の初期状態は , 独自属性のノードのみで構成する . モデル構築時に考えられるすべての属性をノードとして用意しておく , 行動予測の際に用いない属性が存在するためリソースが無駄になる .

そのため , HMM の初期状態ではユーザ独自の属性を表すノードのみを用意し , 他のノードは初めてユーザが訪れた際に新たなノードを動的に構築する . ノードの構築の際に , 内部属性やノード間の状態遷移も構築する .

3.5 本章のまとめ

本研究の目的は、ユーザが訪れたことのない地域に居る際も利用できる行動モデルを構築することである。本研究では隠れマルコフモデルを用いてユーザが滞在した場所の属性をモデル化する MUGI を構築する。MUGI では、ユーザが滞在した場所の属性をノードとして定義し、ノード間の関連性を表現する。それによりユーザの行動における習慣性を抽出する。

次章では、MUGI と MUGI を構築、利用するためのシステムである CHOCO の設計について述べる。

第4章 CHOCOの設計

本章では，訪れたことのない地域でも利用できる行動モデルであるMUGIを動作させるためのCHOCOの設計を述べる．まず，設計方針について述べ，システム構成について述べる．次に，MUGIの三つのモジュールに分けて設計を述べる．

4.1 設計概要

訪れたことのない地域で利用可能な行動モデル MUGI を構築するためのシステムを CHOCO と名付ける。

CHOCO の設計方針は、第 2 章で分類した行動モデルの三要素である、データ収集部、データ分析部、行動モデル構築部は機能ごとのモジュールで構成される。各モジュールは独立に動作しソフトウェアの再利用性を高める。また、ソフトウェア設計において各要素のモジュールが独立に動作していれば、ある部分だけを別のモジュールへ入れ替えられる。例えば、GPS によるデータ収集を、RF タグによるデータ収集に切り替える場合に、他のモジュールへ影響せずに切り替えられる。

4.1.1 ハードウェア構成

CHOCO が動作するハードウェア環境を図 4.1 に示す。位置取得デバイスと携帯計算機は接続されており、位置取得デバイスで取得した位置情報は携帯計算機へ転送される。携帯計算機は、ネットワークへ接続でき、必要に応じてネットワークを介して GIS サーバから情報を取得する。次に、各ハードウェアの要件を示す。

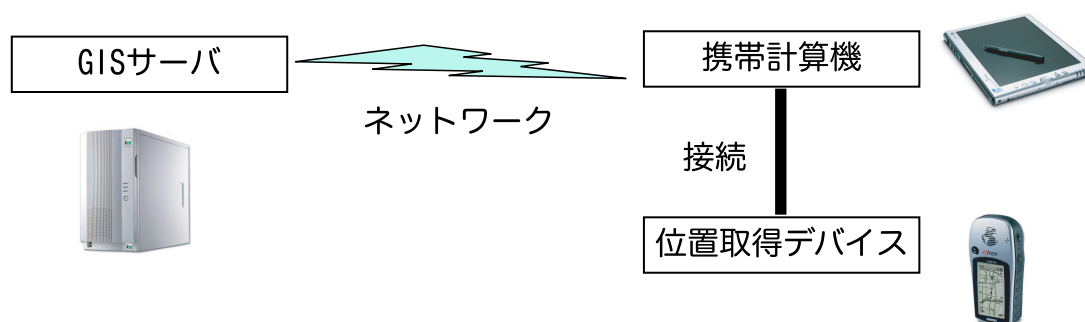


図 4.1: ハードウェア構成

- 位置取得デバイス

ユーザの位置を自分で測位できるデバイス。例えば、GPS 受信機である。RF タグや超音波受信機などのように環境側でユーザの位置を取得する場合は位置取得デバイスは不要になり、ネットワークを介して位置情報を取得する。

- 携帯計算機

CHOCO が動作する計算機資源と、CHOCO を利用したアプリケーションを操作するための GUI を備えた計算機。また、位置取得デバイスからデータを受け取り、GIS サーバへ接続するために、ネットワークへ接続できる必要がある。例えば、タブレット PC や PDA などがある。

- GIS サーバ

地理位置情報を保管し，外部ホストからの地理位置情報要求に対して地理属性を返答するサーバ．

4.1.2 ソフトウェア構成

CHOCO は，行動モデル構築三要素により，ソフトウェアも図 4.2に示すように三つのモジュールに分割して設計する．それらの要素の中には機能ごとにモジュールが存在する．次節にて各モジュールについての詳細を述べる．

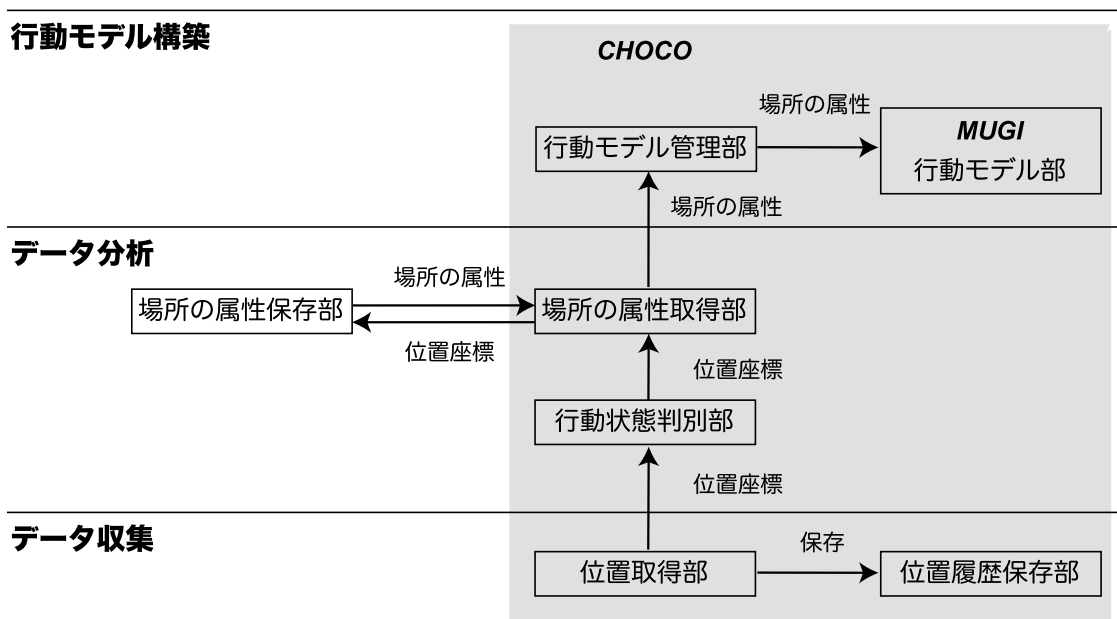


図 4.2: ソフトウェア構成

4.1.3 ソフトウェア動作順序

図 4.3に CHOCO 内での動作順序を UML のアクティビティ図を用いて示す．

1. ユーザの位置情報解析

ソフトウェアの処理開始は位置情報が取得されたと同時に開始し，利用しやすいように位置取得デバイスから取得したデータの構文解析を行う．

2. 行動状態の分析

構文解析された位置情報はデータ分析部へ転送され，ユーザの行動状態を判定する．ユーザの行動状態が滞在状態に変化すると次へ進む．

3. 属性を取得

滞在した場所の属性を取得する。まず、滞在した場所の属性がユーザ独自の属性として定義されているかを調査し、定義されていなければ GIS から滞在している場所の属性を取得する。

4. モデルへ反映

ユーザが行動を開始した場所と滞在している場所の属性を用いてモデルへ反映する。

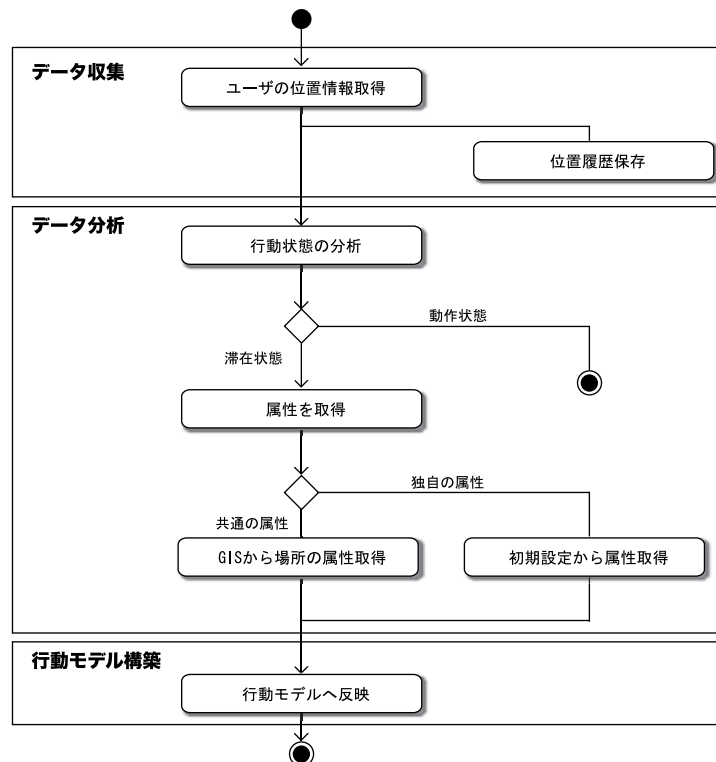


図 4.3: アクティビティ図による CHOCO の動作順序

4.2 データ収集部

ユーザの位置情報を管理する。位置取得デバイスからユーザの位置を検知したセンサデータを解析し、データ分析部へ送信する。

4.2.1 位置取得部

ユーザの位置情報を管理する。位置取得部は、GPS や RF リーダなどの位置取得デバイスから取得したセンサデータの構文を解析し、システム内で利用する位置座標系へ変換

する。

例えば、位置取得デバイスが GPS 受信機である場合、データ分析で利用する測地系が Tokyo であるのに対し、位置取得デバイスの出力する測地系が WGS84 の場合は Tokyo 測地系へ変換する。

4.2.2 位置履歴保存部

過去の位置情報を行動状態の判定に用いるためにユーザの位置情報の履歴を保存する。受信した位置情報をファイルや物理メモリ内などに履歴を一時的に保存する。

本研究で想定している位置取得デバイスからはユーザが停止しているか行動しているかを一度の位置情報データでは判定できないためである。

4.3 データ分析部

ユーザの滞在した場所の属性を取得する。ユーザの位置情報を受信し、ユーザの行動状態を判定する。ユーザが滞在状態に変化したら、その場所の属性を取得し、行動モデル構築部へ送信する。

4.3.1 行動状態判定部

ユーザの行動状態を判定する。ユーザの現在の位置情報を受信し、現在の位置と過去の位置履歴を利用してユーザの行動状態を判定する。行動状態はユーザが滞在しているか、動作しているかの二通りである。

4.3.2 場所の属性取得部

ユーザが滞在している場所の属性を取得する。ユーザが滞在している位置情報を受信し、場所の属性を取得する。

4.3.3 場所の属性保存部

位置と場所の属性を関連づけられた情報が保存する。位置情報を受信し、その位置に関連づけられた場所の属性を返信する。本部分は、ユーザに対して共通の情報を保管する部分なのでネットワークを介して要求を受け付ける。

4.4 行動モデル構築部

行動モデルを構築し、管理を行う。場所の属性を受信し、行動モデルを管理、構築する。

4.4.1 行動モデル管理部

行動モデルの保存，読み込みを行う．行動モデルはユーザの行動の特徴を保存する部分であるため，CHOCOの終了とともに消滅してはならない．そのために，CHOCOが終了した際にも行動モデルはファイルなどに保存し，再びCHOCOが起動した際に行動モデルの状態を復元する．また，行動モデル部へ学習の指令や予測を行う命令を発効する．

4.4.2 行動モデル部：MUGI

MUGIで用いるノードや属性を保管する．行動モデル部が受信する命令は学習命令と，予測命令の2種類である．学習命令を受信した場合は，モデル内のノード間の重みや内部属性の値を更新する．予測命令を受信した場合はモデル内のノード間の重みや内部属性の値を用いて計算し，結果を返す．

4.5 アプリケーションインタフェース

CHOCOがアプリケーションへ提供する機能である．CHOCOへの要件より，次に行動モデルを利用したアプリケーションの例を挙げる．

アプリケーションが次にユーザが訪れる場所を行動モデルを用いて予測し，次の場所への最適な経路を自動で提案したり，訪れる場所の混雑度や定休日などの情報を提示する．

このシナリオを実現するためにCHOCOは行動予測と機能を提供する．次に，機能の詳細を説明する．

行動予測機能は，ユーザがある場所からどの場所へ行くかをMUGIを用いて求める．

行動予測を行うためにはCHOCOに場所の属性，先読み数，結果数を入力する．返り値は，場所の名前と確率の組が入力で与えられた数だけ返却される．表4.1に入出力の一覧を挙げる．

表 4.1: 行動評価機能の入出力

インタフェース	値	説明
入力	場所の属性	現在地となる場所の属性
	先読み数	予測する場所の数
	結果数	確率の高い順に取得する結果数
出力	場所の名前	予測された場所の名前
	確率	場所へ行く確率

4.6 本章のまとめ

本章ではCHOCOのハードウェア、ソフトウェアの設計について述べ、各モジュールの設計を説明した。その上で、アプリケーション例を挙げ、CHOCOが提供する機能を分類した。次章では、CHOCOの実装について述べる。

第5章 CHOCOの実装

前章では，CHOCOの設計について述べた．本章ではCHOCOの実装について述べる．まず，実装環境を述べ，前章において分類した各部の実装を要点となるAPIを抜粋して説明する．

5.1 実装環境

CHOCO を動作させるためには第 4 章の設計より，2 台のコンピュータと位置取得機器が必要である．CHOCO を構成するモジュールはモバイル端末で動作し，場所の属性保存部は GIS サーバで動作する．また，本研究では位置取得機器として GPS を利用する．本機構の実装言語は Java 言語 [17] を用いた．

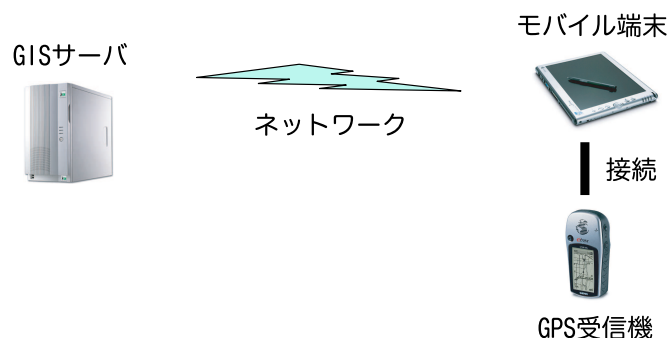


図 5.1: ハードウェア環境

次に，各コンピュータと位置取得機器の仕様を挙げる．

5.1.1 コンピュータ環境

CHOCO を動作させるためには第 4 章の設計より，2 台のコンピュータが必要である．1 台目はユーザがアプリケーションの操作，位置情報の取得のために用いるモバイル端末である．2 台目は地理属性を保存する GIS サーバである．2 台の主な仕様を表 5.1 に掲載する．

表 5.1: 実装環境

項目	モバイル端末	GIS サーバ
CPU	超低電圧版 Intel® Pentium® M 733MHz	Pentium® M 2.00GHz
RAM	512MB	2GB
OS	Windows® XP Tablet PC Edition 2005	Linux kernel 2.4.26 (Vine Linux 3.0)
JDK	Java 2 Platform Standard Edition 5.0	—
Extension	Java Communications API 2.0, JDBC	—
RDBMS ¹	—	PostgreSQL 7.4.5

¹Relational Database Management System

5.1.2 位置取得システム

GPS 受信機は , Garmin 社 [9] の eTrex Vista を用いた . GPS 受信機的主要仕様を表 5.2 に示す .

表 5.2: GPS 受信機仕様

項目	モバイル端末
データ更新時間	1 秒
位置精度	約 15mRMS
出力フォーマット	NMEA0183 , RTCM104(DGPS)

5.1.3 ソフトウェア設計

ソフトウェアの Java 言語におけるパッケージ構成を述べる . CHOCO の各手順は jp.ac.keio.sfc.ht.matsu.mugi 以下で実装する . 以降 , パッケージ名を記載する場合は jp.ac.keio.sfc.ht.matsu.mugi 以下のパッケージ階層を示す . また , 以降の文中ではパッケージ名 , クラス名 , メソッド名は Sans Serif 体を用いて記載する .

表 5.3: 行動モデル三手順とパッケージの対応

手順	パッケージ名
データ収集部	locator
データ分析部	locationanalyzer
行動モデル部	activitymodel

次節以降は , 行動モデル三手順の要素ごとに節を区切り , 4.1.1 節にて分類した機能ごとに分割したモジュールの実装を述べる .

5.2 データ収集部

データ収集部は GPS 受信機から送信される位置情報の構文を解析し , その位置情報が含まれたイベントオブジェクトを発行する . また , データ分析部での行動履歴を分析するために位置の履歴を保存する .

5.2.1 位置取得部

位置取得部は、位置取得デバイスから取得した情報を解析し、抽象化や正規化を行った位置情報をデータ分析部へ送信する。

位置取得部は GPS 受信機から取得した位置情報により処理を開始する。位置取得部は、整形した位置情報を出力する。図 5.2 に位置取得部のクラス図を示す。

SimpleRead クラスは、GPS 受信機から受信した NMEA 0183 フォーマットで位置情報を取得し、NMEAParser を用いて構文解析し、GeoPoint オブジェクトを生成する。GeoPoint クラスは、コンストラクタに緯度経度を与えると、GeoPoint から値を取得する際に任意の測地系へ変換できる。

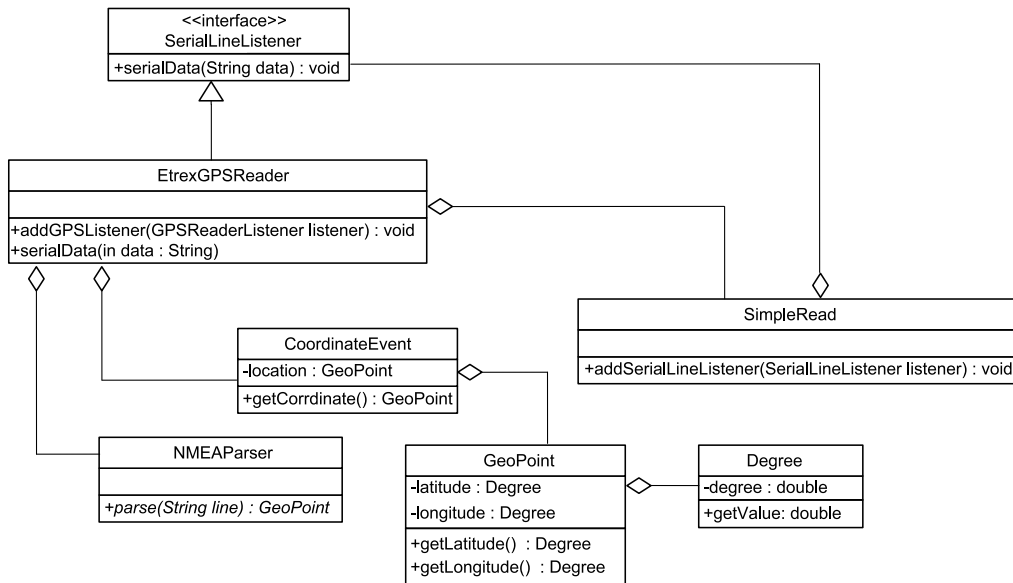


図 5.2: 位置取得部クラス図

5.2.2 位置履歴保存部

位置履歴保存部はデータ分析部で行動常態の判定に用いるために位置履歴位置履歴の保存と読み込みを行う。

履歴を追加するには、GeoPoint オブジェクトを HistoryLogger クラスの `add(GeoPoint location)` メソッドで追加する。履歴を取得するには、`getPrevious()` や `getAllHistory()` メソッドを用いて GeoPoint オブジェクト、または GeoPoint オブジェクトのリストを取得する。図 5.3 に位置履歴保存部のクラス図を示す。

本実装で保存する位置履歴の数は 300 レコードとした。データ分析部では、ユーザの行動常態を判別するために、過去 5 分の位置履歴を用いて判定する。そのため、GPS から位置情報は 1 秒間隔で送信されているので、 $5 \text{分} \times 60 \text{秒} = 300 \text{レコード}$ となる。

本モジュールはデザインパターン [7] の Singleton パターンを用いて実装した。Singleton パターンの利用により、システム内の複数のオブジェクトからこのモジュールを利用できる。

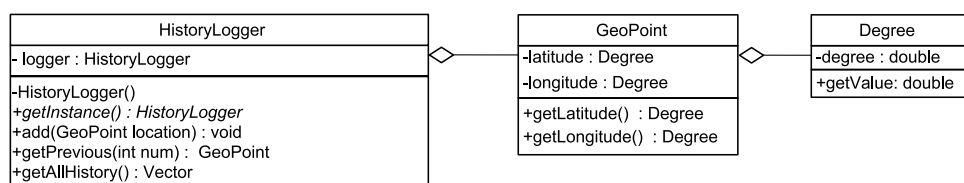


図 5.3: 位置履歴保存部クラス図

5.3 データ分析部

データ分析部は、データ収集部から受信した最新の位置座標と位置履歴を用いて、ユーザの行動状態を判定する。ユーザが滞在した場所の属性と場所の名称をイベントとして発行する。

5.3.1 行動状態判定部

行動状態判定部は JudgeMovement クラスで構成する。動作の流れは、currentPosition() メソッドに現在の位置情報である GeoPoint オブジェクトを引数として与えると、内部で位置履歴を参照し、行動状態のイベントオブジェクトを出力する。図 5.4 に位置取得部のクラス図を示す。

行動状態を求めるために、ユーザの移動速度を利用して判定する。移動速度を求めるためには 2 点の座標と時間がわかれば、距離/時間を計算することで移動速度を算出できる。本実装では、ユーザの位置の遷移から移動速度を求め、ある一定速度以下の時間が経過したら行動状態と判定する。

しかし、実際の GPS 受信機からの位置情報には誤差が含まれており、ユーザが実際は動いていないのに取得した位置情報は移動していることがある。そこで、ユーザの 8 秒間の移動速度の平均を計算し、ユーザの行動状態を判定する。移動速度の平均値を 8 秒間に設定した理由は実際の GPS 受信機の位置情報を用いて平滑化した結果、8 秒が最適であった。平滑化の実験内容は付録 A に記載する。

5.3.2 場所の属性取得部

場所の属性保存部へネットワークを通じて位置情報から最寄りの場所を検索し、その場所の属性を取得する。図 5.6 に場所の属性取得部のクラス図を示す。

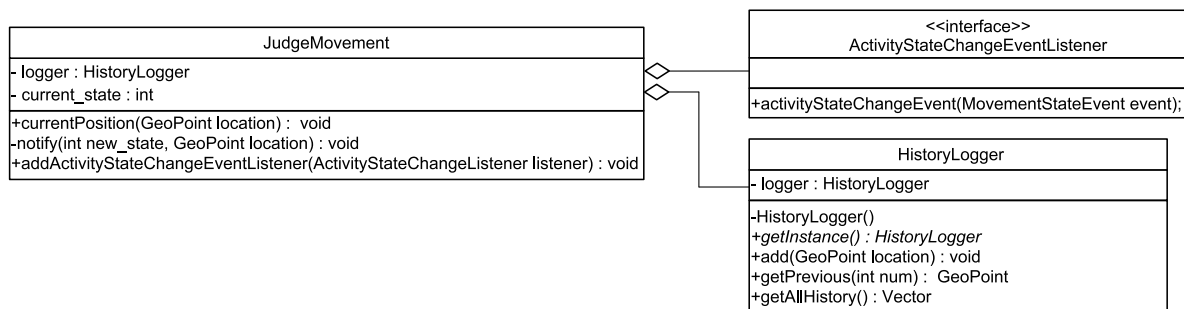


図 5.4: 行動状態判定部クラス図

場所の属性保存部は RDBMS を用いてデータを保管している。プログラムは JDBC を用いて接続し、SQL で問い合わせ、結果を取得する。場所の属性保存部では PostgreSQL を利用し、GIS 情報を保存するために PostgreSQL の拡張ライブラリである PostGIS[10] を用いる。

PostGIS は PostgreSQL が空間座標を保存できるように拡張するライブラリである。また、保存した空間座標への問い合わせを可能にする。例えば、あるテーブル 'thetable' には二つのカラムがある。一つ目は位置座標が保存されている geometry カラム、二つ目は位置の名称が保存されている symbol カラムが定義されている。座標 (100, 100) から距離 50 以内にある名称を取り出したい場合は図 5.5 の SQL で一覧を取得できる。

```

SELECT symbol FROM thetable
WHERE Distance(GeometryFromText('POINT(100 100)', -1), geometry) < 50;
  
```

図 5.5: PostGIS を利用した SQL による問い合わせ例

図 5.5 の WHERE 句で用いている POINT, GeometryFromText, Distance は PostGIS により拡張された関数である。図 5.5 で用いている関数の詳細は、まず POINT 関数により (100, 100) の座標を表現し、GeometryFromText 関数により、仮の測地系である -1 へ変換している。次に、Distance に指定した二つの座標の距離を計算している。

5.3.3 場所の属性保存部

本実装では場所の属性の分類に “MAPPLE デジタルデータ SHAPE 版 縮尺 1/10000 (Rel2)[21]” の関東地方の地図情報を利用した。

本部分は不特定多数のユーザへサービス提供を行うため、専用のサーバで動作する。場所の属性は PostgreSQL のネットワークのネットワーク接続機能を利用し、場所の属性を求めたい場合は、ネットワークを通じて SQL を用いる。

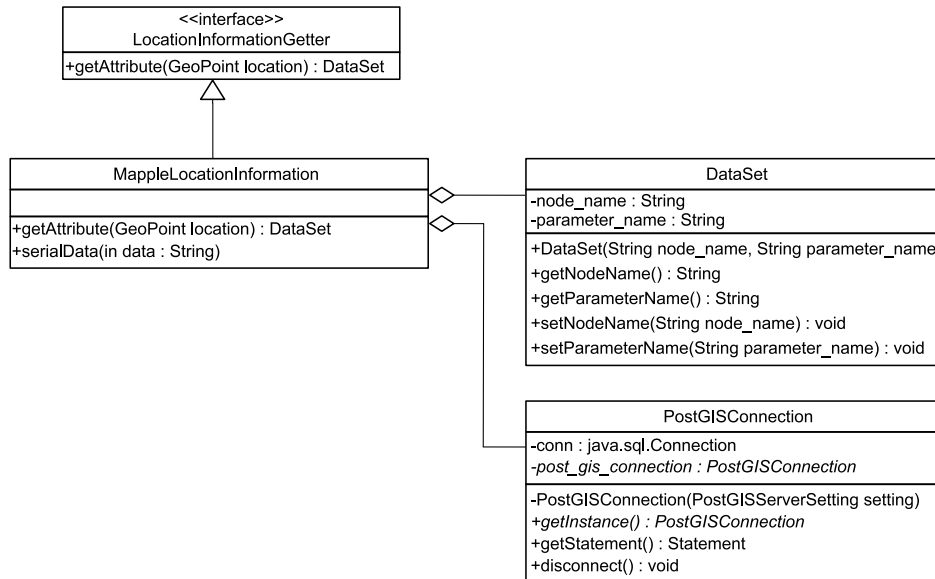


図 5.6: 場所の属性取得部クラス図

5.4 行動モデル構築部

行動モデル構築部では，データ分析部より取得した場所の属性を HMM によりグラフ化する．

5.4.1 行動モデル管理部

行動モデル管理部は行動モデル部の保存，読み込みを行うクラスである．

行動モデル部は，行動履歴より構築した個人の行動の特徴であるため，CHOCO を再起動した場合にも再起動前と同じモデルを復元する必要がある．そのために，行動モデル管理部が行動モデルの常態を指示のあった場合にファイルへ保存，復元を行う．

行動モデル部を管理するクラス名は，HMMManager であり，行動モデル部を読み込む場合は load() メソッドを用い，保存する場合は save() メソッドを用いる．

5.4.2 行動モデル部：MUGI

行動モデル部は HMM を実装した部分である．行動モデル部は MUGI の学習と MUGI からの情報取得機能を提供する．図 5.7 に位置取得部のクラス図を示す．

行動モデル部は HMM，HMMNode の二つのクラスから構成される．HMM の役割は，HMM のノードを表す HMMNode オブジェクトの管理と HMM の学習である．

HMM クラスでは，HMM のノードを管理するために HMM のノード名である String オブジェクトと HMM のノードである HMMNode オブジェクトを組にして Hashtable へ

保存している。また，HMM の学習機能は，内部状態の学習とノード遷移の学習がある。内部状態の学習は，学習させるノード名と内部常態を保持した Dataset オブジェクトを learn(Dataset dataset) メソッドで行い，ノード遷移の学習は，遷移元ノード名と遷移先ノード名を link(String from, String to) メソッドで行う。

HMMNode クラスは HMM の内部状態と，他のノードへの遷移数を保存している。内部状態は，過去に訪問した場所の名前と回数を Hashtable で保存し，他のノードへの遷移数は，リンク先のノード名と回数を Hashtable で保存している。

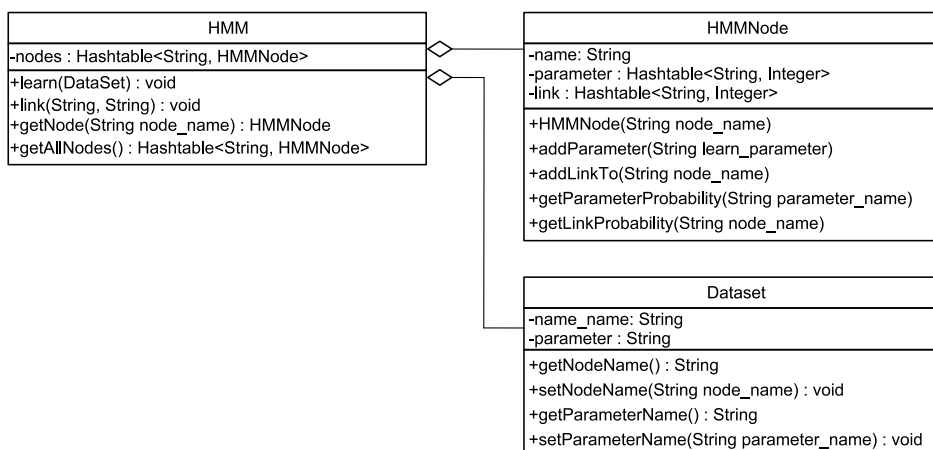


図 5.7: 行動モデル保存部クラス図

5.5 アプリケーションインタフェース

CHOCO がアプリケーションへ提供するインタフェースを記載する。CHOCO はユーザが滞在している場所の属性から指定した回数だけ先の最も高い確率で遷移する場所の名称と属性を返す。

```

/**
 * 行動予測を行う
 * @param String start_place 開始地点の場所の属性
 * @param String steps      予測回数
 */
public Dataset getNextProb(String start_place, int steps);
  
```

CHOCO を利用するためのサンプルコードを図 5.8 に示す。

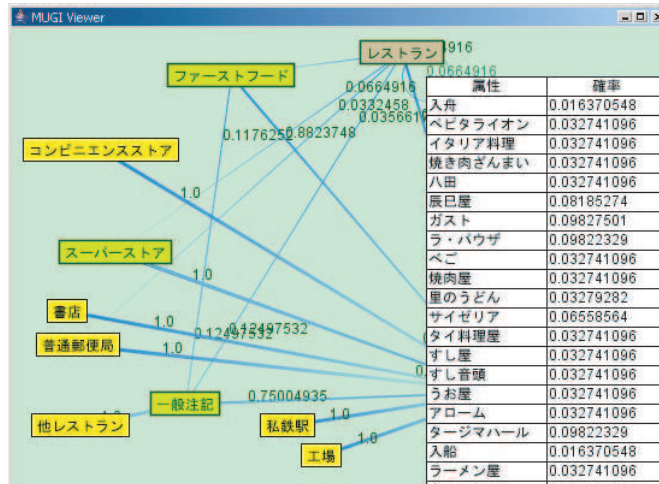


図 5.10: HMM をの内部属性を表示

5.7 本章のまとめ

本章では, CHOCO の実装環境を述べた. 次章では, MUGI 及び CHOCO の定量的, 定性的評価を行う.

第6章 CHOCOの評価

本章では、行動の習慣性を表現する行動モデル、MUGIと、MUGIを構築するためのシステム、CHOCOを評価する。まず、MUGIの定量的評価を行い、CHOCOの定性的評価を行う。

6.1 定量的評価

本節では隠れマルコフモデルを用いた行動モデル MUGI の定量的評価を行う。MUGI は学習した情報から予測を行うため、ユーザが CHOCO を利用しだした時から利用可能である。行動モデルの学習を一定期間行うとよりユーザの習慣性を反映できたモデルを構築できると考える。どの程度の学習期間が必要かを自分の行動履歴を用いて評価する。

6.1.1 学習の収束度

MUGI の学習の収束期間を測定するために自分自身の行動履歴用い、MUGI に学習させる。普段の行動の習慣性を取得するために、行動履歴はあらかじめ自分が普段行動を行う地域のみ履歴を使用する。学習に用いた行動履歴の詳細を表 6.1 に挙げる。また、用いた行動履歴を図 6.1 に掲載する。

項目	内容
日数	62 日分
滞在回数	178 回
1 日の平均移動回数	約 2.871 回
ポイント数	71730 トラック
ファイルサイズ	4.85MB

表 6.1: 用いた行動履歴の詳細



図 6.1: 使用した行動履歴を地図にプロット

次に、表 6.1 に挙げた行動履歴を MUGI へ学習させた際の MUGI の属性値の変化をグラフにする。

表 6.2に滞在数に対する MUGI のノードの数をグラフにした．MUGI においてノード数が表しているのは，過去に滞在した場所の属性の数である．109 滞在目（約 37 日目）で 12 ノードに収束した．約 1ヶ月と少しで習慣的に行動する場所の属性は網羅したことがわかる．また，ノード数を y と置き，滞在回数を x と置くと，式 (6.1) に R2 乗値 0.9516 で近似する．

$$y = 2.3379Ln(x) + 0.284 \quad (6.1)$$

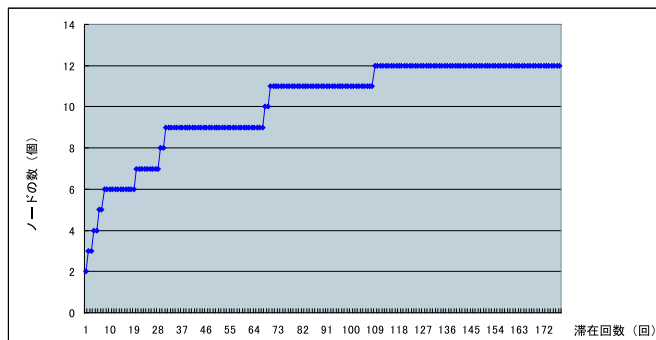


図 6.2: 滞在数に対するノード数の変化

表 6.3に滞在数に対する MUGI のリンク数をグラフにした．表 6.2と表 6.3の結果より，滞在する場所の属性は変化しないが，行動遷移の種類は変化するといえる．また，リンク数を y と置き，滞在回数を x と置くと，式 (6.2) に R2 乗値 0.9834 で近似する．

$$y = 0.2112x + 6.2182 \quad (6.2)$$

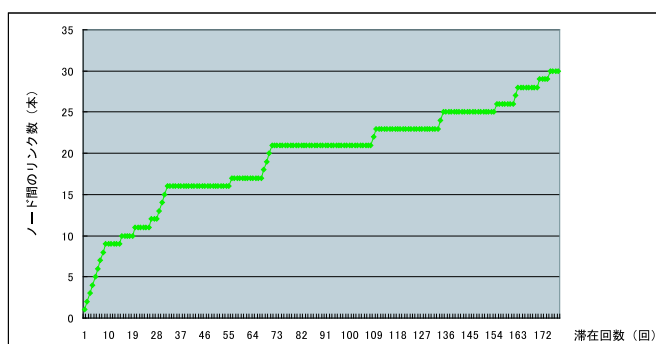


図 6.3: 滞在数に対するリンク数の変化

表 6.4に，滞在数に対する MUGI 内のノードが保持する内部属性数をグラフにした．表 6.2と表 6.4の結果より，滞在する場所の属性は変化しないが，滞在する場所自体は変化す

るといえる．また，パラメータ数を y と置き，滞在回数を x と置くと，式 (6.3) に R2 乗値 0.9121 で近似する．

$$y = 0.1146x + 9.3769 \quad (6.3)$$

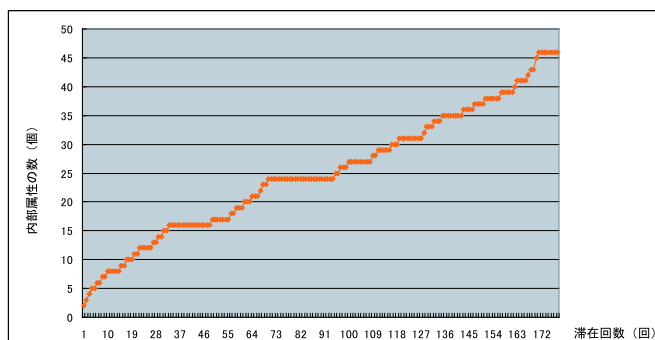


図 6.4: 滞在数に対するパラメータ数の変化

図 6.2 から図 6.4 までの結果より，滞在回数に対してノード数は収束するが，リンク数やパラメータ数は正の傾きで上昇する．これより，ユーザの滞在する場所の属性は限られているといえ，場所の属性が共通した場所へは一定の場所とは限らないといえる．

6.1.2 スケーラビリティ

CHOCO はシステムが終了しても MUGI を永続化するためにファイルシステムへ保存する．CHOCO を使い続けるにつれ，行動履歴数も増加する．よって，行動履歴数とファイル容量の関係を評価する．

本研究では MUGI を Java 言語で実装している．MUGI を保存するために Serializable インタフェースを実装し，オブジェクトを直列化して保存できる．

ファイルサイズを y と置き，滞在回数を x と置くと，式 (6.4) に R2 乗値 0.9563 で近似する．

$$y = 10.56x + 1152.8 \quad (6.4)$$

(6.4) の近似関数に 100 年後のファイル容量を当てはめてみる．100 年後の滞在回数は式 (6.5) より，約 104862 回である．

$$(178 \div 62) \times 365.25 \times 100 = 104862.096774194 \quad (6.5)$$

100 年間の滞在回数 104862 回を式 (6.4) へ代入する

$$f(104862) = 10.56 \times 104862 + 1152.8 = 1108495.52 = \text{約 } 1.1\text{MB} \quad (6.6)$$

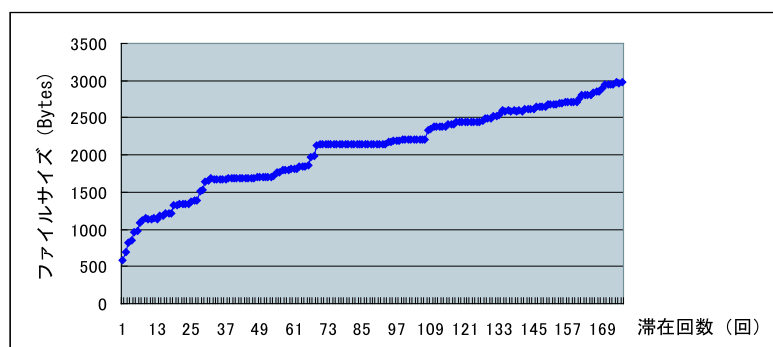


図 6.5: 滞在回数に対するファイルサイズの変化

式 (6.4) より 100 年間使用した場合，CHOCO を保存するファイル容量は約 1.1MB である．これは，数ギガバイトの記憶容量がある PDA やモバイル PC の 1% に満たないため，十分にスケールするといえる．

6.2 定性的評価

本節では CHOCO と既存の関連研究を比較する．既存の行動モデル構築手法である Learning Significant Location and predicting User Movement with GPS[1] と Inferring High Level Behavior from Low Level Sensors[14] を比較する．

本研究では，行動履歴のない地域でも利用できる行動モデルとして既存研究より優位である．

行動モデル構築手法	履歴の無い地域での利用	履歴のある地域での利用	メタ情報の必要性
関連研究 [1]	×		
関連研究 [14]	×		
MUGI			

表 6.2: 用いた行動履歴の詳細

6.3 本章のまとめ

本章では定量評価として MUGI の基本性能を評価し，MUGI の実行速度を評価した．また，定性評価として関連研究と性能面で評価する．次章では今後の課題を述べ，最後に本論文をまとめる．

第7章 結論

本論文では，HMMを用いた行動モデルMUGIを考案し，MUGIの構築・利用を行うためのCHOCOを設計し実装した．本章では今後の課題を挙げ，最後に本論文をまとめる．

7.1 今後の課題

今後の課題は、行動モデルの精度向上と複数の行動モデルの統合について研究を進める。

7.1.1 習慣性の変化に対応

MUGI は人の日常の行動における習慣性を表現した行動モデルである。行動履歴の学習数が増えれば増えるほどモデル内で扱っている確率の母数が増えてゆくため、一つの行動自体の重みが少なくなる。そのため、ユーザに新たな習慣が生まれた場合には学習されづらい。

よって、学習に用いたデータの重みを変化させる必要がある。アプローチとして、過去のデータを学習から除外するか、最近のデータに重みをつける。

7.1.2 フィードバック機構

MUGI を構築するために用いた位置情報はユーザの行動を肯定的に解釈し、行動モデルの構築に用いている。しかし、行動は意に反しない場合もあり、その場合はフィードバックを行うことでモデルに反映できると考える。フィードバックを行うためには、アプリケーションからのフィードバックを受け入れるためのインタフェースを用意する。

また、人間の行動の規則は、時間とともに一定では無いため、過去の行動パターンと最近の行動パターンを同等に扱ってはいけない。過去の行動パターンには重みを軽く、最近の行動パターンには重みを強くする。

7.1.3 行動モデルの共有

本研究では個々人の行動モデルを構築した。未知の地域で、複数人での場合に CHOCO を用いると、アプリケーションが個々人それぞれの行動モデルを利用して提案を行うため、各ユーザへ知らされる情報は別々であるという問題が起る。未知の地域に複数人で居る場合は、複数人の行動履歴を利用し、それぞれのユーザの積集合となる行動モデルを構築しようとする。積集合の行動モデルを構築することで、それぞれのユーザの行動の習慣を抽出した行動モデルを構築できると考える。

7.2 まとめ

本論文では、場所の属性を用いた行動モデル構築手法 MUGI を提案し、MUGI を動作させるための CHOCO を構築した。

謝辞

本研究の機会を与えてくださり，御指導を賜りました慶応義塾大学環境情報学部教授徳田英幸博士に深く感謝致します．また，重要な御助言を頂きました，慶応義塾大学政策・メディア研究科助教授高汐一紀博士の御指導に深く感謝致します．

慶応義塾大学徳田・村井・楠本・中村・南研究会の先輩方には，折りにふれ貴重な助言及び御指導を頂きました．特に村上朝一氏，伊藤昌毅氏には本論文執筆を終えるまでにあたり，絶えざる励ましと御指導を賜りました．また，私の所属している ACE (Active Computing Environment) 研究グループの方々，神武直彦氏，岩井将行氏には貴重な時間を割いて頂いて本研究に関する様々な議論をして頂きました．この場を借りて深く感謝の意を表したいと思います．

平成 17 年 2 月 15 日
松倉 友樹

参考文献

- [1] Daniel Ashbrook and Thad Starner. Learning significant locations and predicting user movement with gps. In *International Symposium on Wearable Computing*, Seattle, WA, October 2002.
- [2] Telecommunications Carriers Association. 社団法人電気通信事業者協会. <http://www.tca.or.jp/>.
- [3] The National Marine Electronics Association. The national marine electronics association. <http://www.nmea.org/>.
- [4] The National Marine Electronics Association. Nmea 0183 standard. <http://www.nmea.org/pub/0183/index.html>.
- [5] U.S. Coast Guard Navigation Center. 2001 gps standard positioning service performance standard final. <http://www.navcen.uscg.gov/>.
- [6] Anind K. Dey and Gregory D. Abowd. Towards a Better Understanding of Context and Context-Awareness. In *Workshop on The What, Who, Where, When, and How of Context-Awareness, as part of the 2000 Conference on Human Factors in Computing Systems (CHI 2000)*, The Hague, The Netherlands, April 2000. ACM Press. Also GVU Technical Report GIT-GVU-99-22.
- [7] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.
- [8] D. Heckerman. A tutorial on learning with bayesian networks, 1995.
- [9] Garmin International Inc. Garmin. <http://www.garmin.com/>.
- [10] Refrations Research Inc. Postgis. <http://postgis.refrations.net/>.
- [11] Changhao Jiang and Peter Steenkiste. *Learning and Recognizing Behavioral Patterns Using Position and Posture of Human*. 12 2.
- [12] Changhao Jiang and Peter Steenkiste. *A Hybrid Location Model with a Computable Location Identifier for Ubiquitous Computing*. Springer, 2003.

- [13] Takamitsu Mizutori. A behavior-based personal controller for autonomous ubiquitous computing. Master's thesis, Keio University, 2004.
- [14] Don Patterson, Lin Liao, Dieter Fox, and Henry Kautz. Inferring high level behavior from low level sensors. In *Proceedings of the Fifth Annual Conference on Ubiquitous Computing (UBICOMP 2003)*, pp. 73–89, 2003.
- [15] L. R. Rabiner and B. H. Juang. An introduction to hidden Markov models. *IEEE ASSP Magazine*, pp. 4–15, January 1986.
- [16] Inc RF Code. Spider tags. <http://www.rfcode.com/>.
- [17] Inc. Sun Microsystems. Java technology. <http://java.sun.com/>.
- [18] Mark Weiser. The computer for the 21st century. In *Scientific American 256(3)*, pp. 94–104, 9 1991.
- [19] 服部可奈子, 小磯貴史, 今崎直樹. 滞在時間を考慮した主要行動パターン抽出方法の検討, 6 2003.
- [20] 国土交通省国土計画局. Gis ホームページ. <http://www.mlit.go.jp/kokudokeikaku/gis/>.
- [21] 昭文社. マップルデジタル地図データシリーズ. <http://www.mapple.co.jp/>.
- [22] 村上知子, 酢山明弘, 折原良平. ベイジアンネットワークを用いた消費者行動モデルの構築実験, 6 2004.
- [23] 北海道大学大学院工学研究科. 積雪寒冷地における震災対応マルチエージェントシステムの開発. http://www.kedm.bosai.go.jp/japanese/kenyukaihatsu/008_sekisettsukanreichi.html.
- [24] 青木茂樹, 大西正輝, 小島篤博, 福永邦雄. Hmm による行動パターンの認識. 電子情報通信学会論文誌 (D-II), vol.J85–D–II, no.7, pp. 1265–1270, 2002.

付録 A 行動状態の判定

ユーザの行動を判定するために用いるアルゴリズムと閾値を求める。ユーザの行動状態の判定に用いる閾値を決定するために GPS 受信機から取得できる値を実測した。

実験の流れを示す

1. 実際に GPS を用いて行動履歴を取得する。
2. 有る一定の平均移動距離を算出する。
3. 滞在している時と動いている時の判別に最もよく当てはまる平均移動距離と閾値を分析する。

実験環境は、実装環境と同じシステムを用いた (p.31)。行動履歴は、531 秒間取得した。その時間内での行動内容を表 A.1 に挙げる。

表 A.1: 実際の移動

経過秒数 (秒)	実際の状態
1-95	徒歩
96-358	滞在
359-513	バイクで移動
514-531	滞在

図 A.1 に示すのは 1, 2, 4, 8, 16, 32, 64 秒間の平均速度をプロットした折れ線グラフである。本測定より 8 秒の平均移動距離を取得し、その平均移動距離が 0.8m/s (時速約 3km/h) を閾値とするのが行動状態の判定に最も当てはまりがよい。

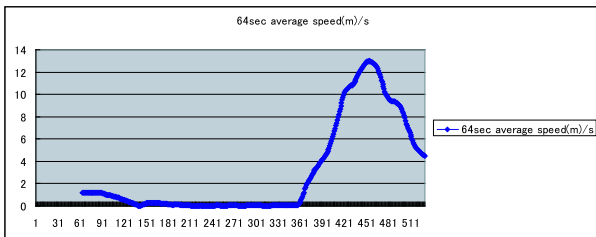
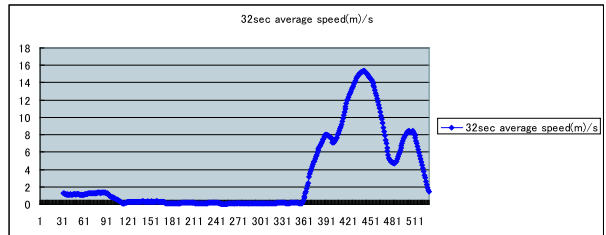
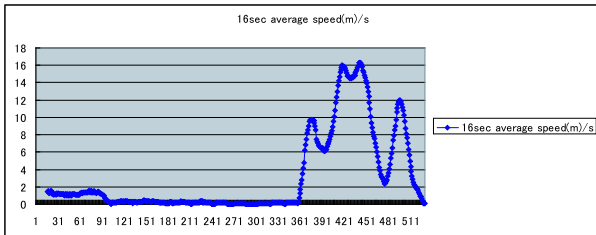
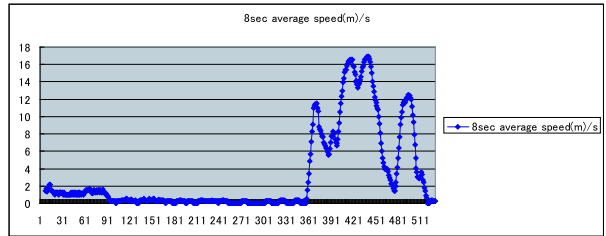
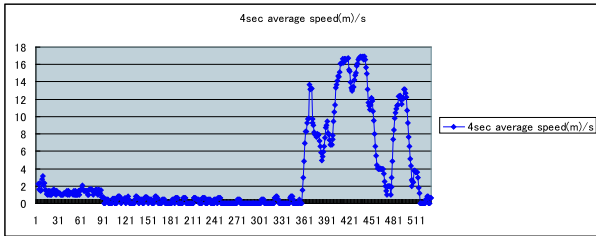
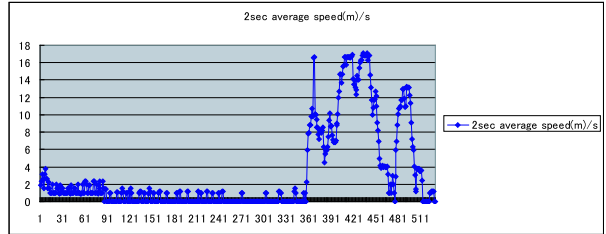
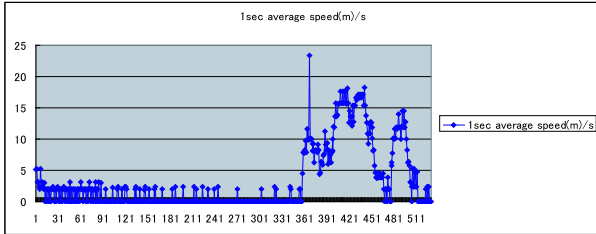


図 A.1: 平均移動距離