

Wi-Fi 位置計測における計測エラーの検出および修正手法

津田 祐輝^{1,a)} 孔 全^{1,b)} 前川 卓也^{1,c)}

概要: 本研究では, Wi-Fi 位置計測における極めて大きい誤差を持つ計測エラーの検出, 修正を行う手法を提案する. Wi-Fi 位置計測において, 参照するアクセスポイントの移動により, 数十 m を超える計測誤差が発生することがある. 提案手法では, 過去の Wi-Fi 計測位置の履歴から現在位置を推定し, Wi-Fi 計測位置と比較することで Wi-Fi 計測エラーを検出, 修正する. このとき, ユーザのコンテキストによって, 位置推定のパフォーマンスが大きく変化するため, 様々な特性を持つ位置推定手法を複数採用することで, コンテキストアウェアなエラー検出フレームワークを構築する. 評価実験では, 実際に様々なコンテキストを想定して, 収集した Wi-Fi 位置計測データを使用し, 提案手法の有効性を検証した.

キーワード: Wi-Fi 位置計測, 計測エラー, モバイル Wi-Fi ルータ, 教師有り機械学習

Detection and Correction of Positioning Error for Wi-Fi based Localization

TSUDA YUKI^{1,a)} KONG QUAN^{1,b)} MAEKAWA TAKUYA^{1,c)}

Abstract: We propose methods that detect and correct very large errors in Wi-Fi based localization. Because of the movement of reference Wi-Fi access points, we sometimes get a Wi-Fi position that has more than dozens of meters errors. In this paper, we detect and correct such errors automatically by comparing current Wi-Fi position with current positions predicted from a history of Wi-Fi positions. However, performances of the position predictors change greatly with the user's context (e.g. migration speed and density of access points around the user). So, we design context-aware frameworks for error detection and correction by employing an ensemble of predictors that have strengths and weakness for different contexts.

Keywords: Wi-Fi based localization, Positioning error, Mobile Wi-Fi router, Supervised machine learning

1. はじめに

近年, 携帯端末を用いた位置計測技術の発展に伴って, 位置情報を用いた様々なサービスが急速に普及している. 現在, 多くの自動車にカーナビゲーションが搭載されており, また高機能携帯端末を介して写真データやマイクロブログの投稿への位置情報の埋め込み, そして SNS 等での位置情報の共有が行われている. 位置計測を行う技術とし

ては, GPS (Global Positioning System) が長年利用されており, 今後も屋外位置計測技術の根幹を担うと考えられる [1]. GPS は複数の GPS 衛星から発信された電波を受信機で受信することで, 現在位置の計測を行っている. GPS 以外の屋外位置計測技術としては, Wi-Fi の無線 LAN アクセスポイント (以下, AP) からの電測情報を用いて位置を計測する手法である Wi-Fi 位置計測が注目されている [2], [3], [4]. Wi-Fi 位置計測では, 携帯端末が受信した Wi-Fi の電波強度とその MAC アドレスを用いて, Wi-Fi 位置情報サービスのプロバイダへ問い合わせを行うことで, プロバイダにより推定された位置情報を得る. そのため, 携帯電話以外にもノート PC, 携帯ゲーム機, 音楽プレイヤー, デジタルカメラ等の無線 LAN 機能を持つ機器

¹ 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

a) tsuda.yuuki@ist.osaka-u.ac.jp

b) kong.quan@ist.osaka-u.ac.jp

c) maekawa@ist.osaka-u.ac.jp

での利用が可能である。

この2手法にはそれぞれ長所と短所が存在する。GPSは、GPS電波の届かない屋内や地下では位置計測を行えない。また、多くの電力を消費するため、電池容量に制限のある携帯端末においては、長時間の利用ができないという問題がある。しかし、周囲が開けた場所であれば、その誤差は数m程度と極めて精度の高い位置計測が可能である。一方、Wi-Fi位置計測は、APが周囲に存在しない場所や利用するWi-Fi位置情報サービスがカバーしていない範囲では位置計測が行えないが、周囲にAPが存在すれば、屋内においても位置計測を行うことができる。また、その消費電力はGPSと比べて少なく、携帯端末での長時間の使用に適している。しかし、その計測誤差は数十m程度とGPSに比べると大きく、計測精度は劣っている。

GPSとWi-Fi位置計測における、このような特長の違いから、多くの研究者によりGPSとWi-Fi位置計測のハイブリッド計測手法が提案されている[5], [6]。これらの手法では、それぞれの長所である高い計測精度や少消費電力を同時に達成することを目指している。また、それぞれの位置計測のカバーエリアを補い合うことが可能である。

一方、近年Wi-Fi位置計測において、数百m以上という極めて大きな計測誤差の発生が増加している。このような計測誤差をもつ位置計測（以下、計測エラー）は、主にAPの移動によって引き起こされている[7], [8], [9]。Wi-Fi位置情報サービスへ登録されているAPの位置と、実際のAPの位置が異なる場合、そのAPを観測した端末の位置は、そのAPが登録されている位置として誤って計測される。このような状況は、住居やオフィスの引っ越しやモバイルWi-Fiルータの存在により発生する。今後、モバイルWi-Fiルータがより普及すると予想されるため、計測エラーへの対応が急がれる。

そこで、本研究ではWi-Fi位置計測における大きな計測エラーの自動検出と計測位置の自動修正を目指す。これらの実現により、Wi-Fi位置計測において計測エラーが検出された際に位置計測をGPSへと切り替えるハイブリッド位置計測が可能となり、またGPSが使用できない状況や端末では、その計測位置を修正できる。

本研究では、計測エラーの検出を外れ値検出とみなす。過去のWi-Fi計測位置やその過程で得られたデータから現在位置を推定する位置推定器を用意し、その推定位置と現在のWi-Fi計測位置を比較することにより、外れ値を検出する。しかし、携帯端末の計測を行った環境（APの密度や端末の移動速度等）が位置推定器のパフォーマンスに多大な影響を与えるため、単一の位置推定器により正確に現在位置を推定することは困難である。そこで、様々な特徴を持つ位置推定器を複数用意し、計測環境を考慮して組み合わせることで計測エラーの検出を行う。また、同様の位置推定器を用いることで、計測位置を修正する。

2. Wi-Fi計測エラーの検出および修正手法

2.1 システムの概要

提案システムの概要は図1となる。提案システムは、「計測エラーの検出」と「計測位置の修正」という2つの目的にそれぞれ対応した機構から成り立っている。それぞれの機構では、教師有り機械学習を用いることで、計測エラーの検出と計測位置の修正を行う。「計測エラーの検出」に対応する機構の出力は、対象とする時刻における計測エラーの有無であり、「計測位置の修正」に対応する機構の出力は、対象とする時刻における修正された現在位置座標である。両機構とも下記のデータを入力とする。

- 時刻 $1 \sim T$ の Wi-Fi 位置計測の現在位置 $(x_1, y_1), \dots, (x_T, y_T)$
- 時刻 $1 \sim T$ の Wi-Fi スキャン WS_1, \dots, WS_T

システムへの入力として、携帯端末がWi-Fi位置計測を実行した際に得られるデータのみを使用することを想定する。これは、Wi-Fi位置計測の利点である短時間計測、低消費電力を損なわないために、追加的なデータの収集（加速度データ等）を行うべきでないと考えるためである。携帯端末が時刻 T にWi-Fi位置計測を実行した際に、計測位置 (x_T, y_T) と問い合わせに用いたWi-Fiスキャン WS_T の2種類のデータが得られる。Wi-Fiスキャン WS_T は、携帯端末が観測した周囲のAPのMACアドレスと受信信号強度からなるデータである。また、携帯端末がWi-Fi位置計測を継続的に利用する場合、過去のWi-Fi位置計測によるデータも保持していると考えられる。したがって、時刻 $1, \dots, T$ におけるWi-Fi位置計測とWi-Fiスキャンの履歴も利用可能なデータとなる。

また、両機構は教師有り機械学習を用いるため教師データとして、時刻 T の正確な現在位置座標 (X_T, Y_T) が必要となる。本研究では、正確な現在位置座標としてGPSによる位置座標とESM (Experience sampling method) によって得られた位置座標を使用する。Wi-Fi位置計測時に屋内にいる等の理由でGPSによる位置計測が行えなかった場合に、実験参加者へ現在位置座標の入力を求めることで、正確な位置を欠かさず取得している。

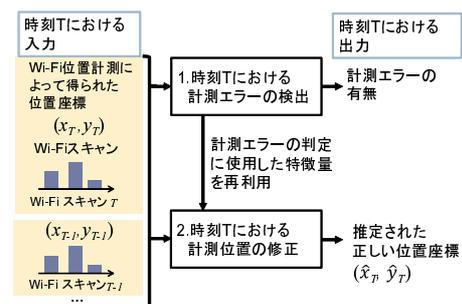


図1 提案システムの概要

Fig. 1 Overview of our proposed method.

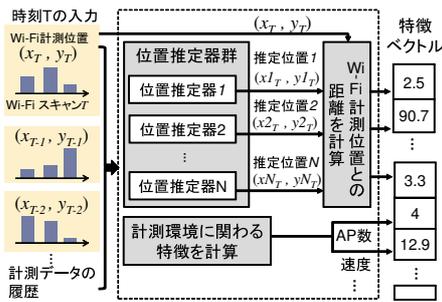


図 2 計測エラーの検出に用いる特徴ベクトルの作成

Fig. 2 Constructing feature vector for error detection.

2.2 計測エラーの検出

計測エラーの検出においては、2クラス分類器を使用して「計測エラー有」もしくは「計測エラー無」の2クラスに分類するモデルを教師有り機械学習により学習する。モデルの入力としては、図2のように特徴抽出を行って構築した特徴ベクトルを使用する。この特徴抽出では、大きく分けて以下の2種類の特徴量を抽出する。

- 位置推定器による推定位置と Wi-Fi 計測位置との距離: 位置推定器 n による推定位置 (x_{nT}, y_{nT}) と Wi-Fi 計測位置 (x_T, y_T) とのユークリッド距離を特徴量として使用する。図2に示すように、位置推定器 n は前述の入力データから位置座標 (x_{nT}, y_{nT}) を推定する。その後、位置推定器による推定位置 (x_{nT}, y_{nT}) と Wi-Fi 計測位置 (x_T, y_T) 間の距離を算出する。この距離が小さければ、位置推定器による予測と観測が近く、計測エラー発生の可能性が低いことを意味する。
- 計測環境の特徴: 計測環境の特徴は、例えば問い合わせに用いた AP の数や携帯端末の移動速度といった、Wi-Fi 位置計測のパフォーマンスや各位置推定器のパフォーマンスに影響を与えるような特徴量である。

2.2.1 位置推定器の詳細

本研究では、計測環境に応じた計測エラーの検出を行うために、以下の様々な特徴を持つ位置推定器を使用する。

カルマンフィルタ: 線形的に表現できる移動軌跡をトラッキングする用途で利用されている技術であるカルマンフィルタ [10] を使用して位置推定を行う。そのアルゴリズムは、予測と更新の2つのステップからなる。予測においては、携帯端末は1つ前の時刻 $T-1$ と同じ速度で移動すると仮定して、時刻 T の位置座標 (x_{pT}, y_{pT}) を予測する。更新においては、時刻 T において予測された位置 (x_{pT}, y_{pT}) を、時刻 T における Wi-Fi 計測位置 (x_T, y_T) をもとに修正・更新し、位置座標 (x_{cT}, y_{cT}) を得る。そして、現在時刻 T の Wi-Fi 計測位置 (x_T, y_T) の影響を受けていない予測位置 (x_{pT}, y_{pT}) を、この位置推定器の時刻 T における推定位置とする。

この位置推定器は、Wi-Fi 計測位置の過去の履歴が十分に存在し、また移動速度の変化が小さい場合に、良いパ

フォーマンスを得る。しかし、Wi-Fi 位置計測の履歴が少ない場合は、過去の軌跡の情報が正確に得られずに位置推定の精度が低下してしまう。また、過去の履歴が最低2つ以上存在しなければ、この位置推定器は位置推定を行うことができない。この場合、対応する特徴ベクトルの要素を欠損値として扱う。

改良カルマンフィルタ: この位置推定器は、前述のカルマンフィルタを改良したフィルタを用いて位置推定を行う。本研究が対象とする入力データは、計測エラーによる極めて大きな誤差を含むことがある。カルマンフィルタは、予測と更新のステップを毎入力データごとに繰り返すため、1つの大きな誤差がその後の予測に重大な影響を及ぼしてしまう。そこで、本研究では時刻 T の Wi-Fi 計測位置 (x_T, y_T) と時刻 $T-1$ の更新位置 (x_{cT-1}, y_{cT-1}) から求めた速度が時速 100km を超過した場合は、Wi-Fi 計測位置 (x_T, y_T) を用いて更新を行わずに更新ステップを終える。このように対応することで、より精度の高い位置推定を行うことができる。

パーティクルフィルタ: 携帯端末の利用者による実際の移動には、車や電車を乗降するといった移動手段の変更や交通信号機等による移動速度の急激な変化が起こることが考えられる。そのような場合に対応するため、非線形な状態遷移を行うシステムの状態を予測するパーティクルフィルタ [11] を使用する。そのアルゴリズムは、サンプリング、重み計算、リサンプリングの3ステップからなる。サンプリングにおいては、1つ前の時刻 $T-1$ のパーティクルから、新たなパーティクルを作成し、移動モデルに基づいてそのパーティクルを移動させる。重み計算においては、パーティクルへの重みづけを行う。本研究では、Wi-Fi 計測位置 (x_T, y_T) に近いパーティクルは、実際にその場に携帯端末が位置する可能性が高いと考えられるため、近いパーティクルに大きい重みを与える。リサンプリングにおいては、重みが小さいパーティクルの排除を行う。この3ステップを繰り返すことで軌跡の推定を行う。この位置推定器の時刻 T における出力としては、サンプリングのステップで得られたパーティクルの平均位置を使用する。

この位置推定器は、Wi-Fi 計測位置の過去の履歴が十分に存在する場合に良いパフォーマンスを得る。ただし、パーティクルの速度を求めるために直前の履歴が2つ必要であるため、過去の履歴が最低2つ以上存在しなければ、この位置推定器は位置推定を行うことができない。この場合、対応する特徴ベクトルの要素を欠損値として扱う。

他プロバイダへの問い合わせ: この位置推定器は、時刻 T の Wi-Fi スキャン W_S_T を用いて、他の Wi-Fi 位置情報サービスのプロバイダへ問い合わせることにより、推定位置を得る。この位置推定器は、Wi-Fi 計測位置の過去の履歴が存在しない場合でも位置推定が可能であるが、 W_S_T に含まれる AP の数が少ない場合は正確な位置推定が行え

ない場合がある．さらに Wi-Fi 位置計測を行うため，異常 AP の影響を受けることが考えられる．しかし，Wi-Fi 位置情報サービスのプロバイダによって，AP の登録方法や登録状況が異なるため，各プロバイダのカバーする範囲や位置推定の精度に違いが生じる．そのため，他のプロバイダの問い合わせから得られた推定位置により，対象とするプロバイダの Wi-Fi 計測エラーの検出を行うことができる．

修正 Wi-Fi スキャンによる問い合わせ: この位置推定器は，異常 AP を Wi-Fi スキャン WS_T から取り除き，修正後の Wi-Fi スキャン WS'_T を用いて，入力データの Wi-Fi 位置計測に用いたプロバイダへ再度問い合わせを行うことで位置推定を行う．以下に，具体的な例を用いて，この位置推定器が行う位置推定の 4 つのステップを説明する．ここでは，説明のために Wi-Fi スキャン WS_T が 6 つの AP (AP1, ..., AP6) からの信号を含み，そのうち 1 つが異常 AP (AP1) であるとする．

- (1) Wi-Fi スキャン WS_T に含まれる AP を g 個のグループにランダムに分割する．このとき，それぞれのグループ内の AP の数は， $\left\lceil \frac{\# \text{APs in } WS_T}{g} \right\rceil$ を超えないものとする．また， g は奇数である必要があり，ここでは $g = 3$ として説明を行う．例においては，グループ 1 に AP1 と AP2，グループ 2 に AP3 と AP4，グループ 3 に AP5 と AP6 というように分割される．
- (2) それぞれのグループ内の AP の信号強度情報から構成される Wi-Fi スキャンを作成し，Wi-Fi 位置情報サービスに問い合わせを行う．そして， g 個の Wi-Fi 計測位置を得る．この例においては，3 つのグループに対応する Wi-Fi スキャンが作成され，Wi-Fi 計測位置 (x_{g1T}, y_{g1T}) , (x_{g2T}, y_{g2T}) , (x_{g3T}, y_{g3T}) を得る．
- (3) 得られた Wi-Fi 計測位置を比較し，異常 AP を含むグループを特定する．例において，グループ 1 は異常 AP を含むため，対応する Wi-Fi 計測位置 (x_{g1T}, y_{g1T}) は，大きな誤差を持つはずである．そのため，Wi-Fi 計測位置 (x_{g1T}, y_{g1T}) は，他の 2 つの Wi-Fi 計測位置から遠い位置に存在すると考えられる．そこで，離れた計測位置を自動的に特定するために，得られた g 個の Wi-Fi 計測位置に対して階層的クラスタリングを行う．そして，最も大きいクラスタ内の要素の数が， $\left\lfloor \frac{g}{2} \right\rfloor$ を超えたとき，クラスタリングを終了する．得られた最大のクラスタは，異常 AP の影響を受けた Wi-Fi 計測位置を含まず，携帯端末の実際の位置に近い Wi-Fi 計測位置のみで構成されることが期待される．例においては，グループ 2 とグループ 3 に対応する Wi-Fi 計測位置 (x_{g2T}, y_{g2T}) , (x_{g3T}, y_{g3T}) が最大のクラスタを構成することとなる．
- (4) 最大のクラスタ内の Wi-Fi 計測位置を得る問い合わせに用いた AP 全てを含む Wi-Fi スキャン WS'_T を作成する．そして，Wi-Fi 位置情報サービスへ問

合わせを行い，推定位置を得る．この例では AP3, AP4, AP5, AP6 を用いて，Wi-Fi スキャン WS'_T を作成する．そして，問い合わせを行うことで異常 AP (AP1) の影響を受けていない Wi-Fi 計測位置を得る．

ここで，多くの Wi-Fi 位置情報サービスは，1 つの AP のみから成る Wi-Fi スキャンによる問い合わせを受け付けないという問題がある．そのため，2 つ目のステップにおいて 1 つの AP からなるグループが発生する場合，AP 数が 2 となるまで Wi-Fi スキャン WS_T からランダムに AP を選択してグループに追加する．本研究では，AP を分割するグループ数 g が 3 の場合と 5 の場合でそれぞれ位置推定器を用意する．

2.2.2 計測環境の特徴の詳細

本研究では，携帯端末の計測環境の情報をとらえるため，以下の特徴量を使用する．

計測間隔: この特徴は時刻 T と前回の計測時刻 $T-1$ の時間間隔に対応する．トンネル内を移動する場合など，周囲に AP が存在しない状況では，長時間 Wi-Fi 位置計測が行えずに計測間隔は大きくなり，軌跡を用いる位置推定器のパフォーマンスが下がってしまう．この特徴量を使用することで，そのようなパフォーマンスの低下を考慮することが可能となる．

移動速度と移動距離: 本研究では，加速度センサ等の他のセンサを用いないことを前提とするため，携帯端末の移動速度は，時刻 $T-1, T$ の Wi-Fi 計測位置 (x_{T-1}, y_{T-1}) , (x_T, y_T) より算出を行う．また，その 2 地点間の距離も特徴量として使用する．この 2 つの値が大きければ，時刻 T において計測エラーが発生していることが予想される．

移動速度の分散と平均: 移動速度の分散は，過去 n 分間の Wi-Fi 計測位置 (x_{T-n}, y_{T-n}) , ..., (x_T, y_T) をもとに算出される．携帯端末の移動速度が変化した場合，この値は大きくなり，同じ速度の移動を前提とするカルマンフィルタを用いた位置推定器のパフォーマンスは低下する．また，移動速度の平均も同様に算出して特徴量とする．

AP 数: 時刻 T における Wi-Fi スキャン WS_T 内に含まれる AP の数を特徴量とする．AP 数が少ない場合は，Wi-Fi 位置計測自体の精度が低くなるため，再度 Wi-Fi 位置計測を使用する位置推定器のパフォーマンスは低下する．

受信信号強度の平均と分散: 時刻 T の Wi-Fi スキャン WS_T 内に含まれる AP の受信信号強度の平均と分散を，特徴量として使用する．これらの特徴は，周囲の電波環境に影響されるため，Wi-Fi 位置計測を使用する位置推定器のパフォーマンスに影響を与える可能性がある．

モバイル Wi-Fi ルータに関する特徴: 携帯端末の周囲に存在するモバイル Wi-Fi ルータは，Wi-Fi 位置計測に悪影響を及ぼす．例えば，モバイル Wi-Fi ルータを所持している人物が，同じ車両に乗車している場合，携帯端末は移動しているにもかかわらず，Wi-Fi 位置計測の結果はそのモ

パイル Wi-Fi ルータがプロバイダに登録されている場所を指し続けることがある。このような場合、携帯端末はモバイル Wi-Fi ルータからの信号を受信し続けることになる。しかし、家庭やオフィスに設置された静的な AP に関しては、携帯端末が高速で移動し続けるため、連続的に観測されることは少ない。すなわち、時刻 $T-1$ に観測された静的 AP は、時刻 T においては観測されることは少ない。そこで、モバイル Wi-Fi ルータに関する特徴として、時刻 $T-1$ と T のいずれかで観測された AP の数に対して、時刻 $T-1$ と T の両方で観測された AP の数の割合を算出して使用する。

さらに、過去 n 分間の Wi-Fi スキャン WS_{T-n}, \dots, WS_T をもとに算出する特徴も使用する。まず、過去 n 分間に 1 度でも観測された AP に関して、過去 n 分間の Wi-Fi スキャン WS_{T-n}, \dots, WS_T で観測された回数を数える。そして、その出現回数の分散を特徴量として用いる。このとき、モバイル Wi-Fi ルータによる AP の観測回数は大きな値を取るが、静的 AP の観測回数は 1,2 回となる。すなわち、モバイル Wi-Fi ルータの所持者が、同じバスや車両に乗りしている場合、その出現回数の分散は増加する。この 2 つの特徴量を用いることで、モバイル Wi-Fi ルータの存在をある程度、検知することができる。

本研究では、過去のどれだけのデータを参照して特徴計算するかを決める n の値は 5 分とする。

2.3 Wi-Fi 位置計測における計測位置の修正

図 3 は、計測位置の修正機構の全体像を示す。ここでは、計算コストの削減のため位置推定器の推定位置と計測環境の特徴量を再利用する。前述のとおり、各位置推定器のパフォーマンスは計測環境に応じて変化するため、各推定位置を一律に考慮して、計測位置の修正を行うべきではない。そこで、計測環境の特徴量から時刻 T の計測環境における各推定位置の誤差を推定して、その値を考慮することで、計測位置の修正を行う。各推定位置の誤差は、図 3 の誤差推定器において時刻 T の計測環境の特徴量から、回帰分析を用いて、時刻 T における各推定位置の誤差 e_{nT} を推定する。このとき、位置推定器の誤差とは GPS や ESM によって得られた実際の携帯端末の位置と位置推定器の推定位置間の距離である。このようにして推定された推定位置の誤差 e_{nT} を用いて、式 1 に示す重み付け平均位置 (\hat{x}_T, \hat{y}_T) の算出を行い、修正された計測位置として出力する。

$$(\hat{x}_T, \hat{y}_T) = \frac{\sum_{i=1}^N 1/e_{iT}(x_{iT}, y_{iT})}{\sum_{n=1}^N 1/e_{nT}} \quad (1)$$

式 1 では、推定された誤差 e_{nT} の逆数を、推定位置 (x_{nT}, y_{nT}) の信頼性を表す係数として用いている。推定誤差 e_{nT} の逆数を係数として使用することで、携帯端末の実際の位置に近いと推定された推定位置をより重く考慮する

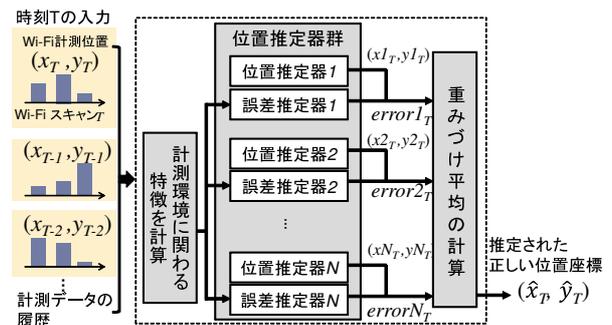


図 3 計測位置の修正機構の全体像

Fig. 3 Overview of correct coordinate prediction.

ことができる。

3. 評価実験

3.1 実験対象

実験対象データの収集のため、複数人の参加者にデータ収集用の携帯端末 (Google Nexus One) を持ち歩いてもらった。そして、携帯端末上のデータ収集用のアプリケーションを使用して、データの収集を行った。Wi-Fi 位置情報サービスのプロバイダは、Google Location Service[12] を使用し、1 分ごとに位置計測を行った。本研究の有効性を評価するためには、様々な計測環境に対応したデータを用意する必要がある。そこで、以下の 3 種類の条件において、移動データを収集した。収集したそれぞれの軌跡データは、始点と終点が全て異なるものである。

- 通勤通学データセット: このデータセットは、参加者の通勤通学時にデータ収集用の携帯端末を持ち歩いてもらうことで収集した。移動方法は電車とバス、徒歩の組み合わせであり、速い速度での移動時や移動速度の急激な変化を想定している。
- ビル街データセット: 参加者がデータ収集用の携帯端末を持って、ビル街を徒歩で散策することにより、移動データを収集した。参加者は、ビル間の歩道や商業ビル、駅建物内を歩くため、頻繁な移動方向の変更や、屋内屋外への移動を想定している。
- 地下街データセット: 参加者がデータ収集用の携帯端末を持って、地下街を徒歩で移動することにより、移動データを収集した。このデータセットは、周囲の AP が少数であるような環境を想定している。

表 1 は、これらのデータセットの概要を示している。また、図 4 は Wi-Fi 計測位置の誤差の分布を、図 5 はその誤差の CDF (累積分布関数) を示している。ここで、誤差とは Wi-Fi 計測位置と GPS や ESM による計測位置間の距離を表している。図 4 より、0~150m には多数の誤差がみられるが、150m 以上では誤差の発生がまばらである。しかし、図 5 から数十万 m 程度の誤差も発生しており、広い範囲で誤差が発生していることが分かる。以上から、150m 以上の誤差を持つ Wi-Fi 計測位置を計測エラーとした。

表 1 各データセットの概要
Table 1 Overview of our data set.

| | 通勤通学 | ビル街 | 地下街 | 全体 |
|---------------|-------|------|-------|-------|
| 軌跡数 | 11 | 10 | 10 | 31 |
| 平均計測時間 (分) | 84.8 | 31.6 | 26.7 | 48.9 |
| 位置計測数 /1 時間 | 24.8 | 44.8 | 39.1 | 31.5 |
| エラー率 (> 150m) | 32.8% | 9.8% | 20.3% | 22.4% |

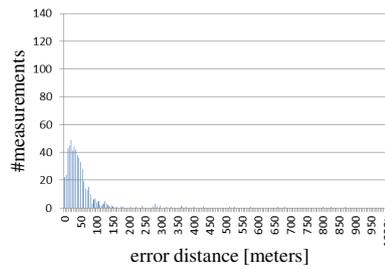


図 4 Wi-Fi 計測位置の誤差の分布

Fig. 4 Distribution of Wi-Fi positioning errors.

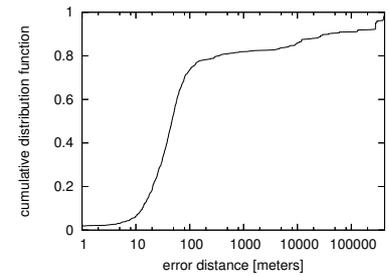


図 5 Wi-Fi 計測位置の誤差の CDF

Fig. 5 CDF of Wi-Fi positioning errors.

4. 実験結果

4.1 Wi-Fi 位置計測における計測エラーの検出

評価手法: 評価のために、表 1 の実験データに対して、leave-one-out 交差検定を行った。これにより、全 31 個のデータのうち 1 つのデータをテストデータとし、残る 30 個のデータをモデルの学習データとして用いることを繰り返すことで、全ての移動データに対して検定を行った。計測エラーの検出を行う 2 クラス分類器としては、C4.5 決定木を用いた。提案手法の有効性を評価するための比較手法として、前述の位置推定器それぞれによる推定位置 (x_{nT}, y_{nT}) と Wi-Fi 計測位置 (x_T, y_T) 間の距離のみを特徴量として構築した 2 クラス分類器 (C4.5 決定木) を用いる手法を用意した。評価項目としては、各時刻における「計測エラー有」「計測エラー無」という 2 クラスへの分類結果の適合率、再現率、F 値および、2 クラスの F 値の平均である平均 F 値を使用する。また、他のプロバイダへの問い合わせを用いる位置推定器においては、「Skyhook wireless」[13] への問い合わせにより位置推定を行った。

評価結果・考察: 図 6 に各手法による計測エラー検出の平均 F 値を示す。表 1 にあるように「計測エラー無」のサンプル数に対して、「計測エラー有」のサンプル数が少ないため、「計測エラー有」のサンプルに重みづけ学習 (LWL:Locally Weighted Learning) を行った結果も図 6 に示している。図 7 には、各手法における計測エラー発生時の分類結果の平均 F 値を各データセットに対して算出したものを示している。表 2 には、各手法の計測エラー発生を分類した詳細な結果を示しており、それぞれの手法において、分類した 2 クラスとその平均の適合率、再現率、F 値を示している。以下、これらの図表から、それぞれの手法の結果について論じる。

まず、過去の軌跡を用いる 3 つの位置推定器の結果に注目する。これらの中では、パーティクルフィルタの結果が最も良い。これは、参加者の移動速度や移動方向の頻繁な変化に対応し、計測エラーの発生を検出できたためである。しかし、通勤通学、地下街データセットにおいては、次のような状況が頻繁に現れており、位置推定の精度を低下させ

ていることが分かった。参加者がトンネルを通る場合や周囲に店の少ない地下通路を移動する場合は、参加者の周囲に AP が存在せず、長時間 Wi-Fi 位置計測を行えない。その場合は、Wi-Fi 位置計測が途切れる前の軌跡をもとに現在位置を推定するため、実際の位置と推定位置間の距離が数百 m 程度と大きくなってしまふ。その結果、計測エラーとして誤検出してしまったと考えられる。これは、他の 2 つの位置推定器についても言えることである。対して、改良カルマンフィルタに注目すると、地下街データセットにおいてパーティクルフィルタの精度を上回っていることがわかる。これは、移動した地下通路は直線的に設計されることが多いため、移動データが直線的な移動を前提とする改良カルマンフィルタに適合したためだと考えられる。

次に、他のプロバイダへの問い合わせによる結果に注目する。図 6 から、この手法は他の位置推定器を用いた手法と比べて、良い結果が得られていることがわかる。しかし、この手法は参加者が大学や公園や駅といった広い面積のエリアを移動するとき、正確な位置推定を行うことができない。そのため、通勤通学、ビル街データセットにおける精度が他の手法と比べると比較的低下している。これは、Wi-Fi 位置情報サービスのプロバイダとして使用した Skyhook Wireless が war-driving[14] によって、Wi-Fi 位置情報データを収集しているために、道路から離れた位置の推定があまり正確でないためと考える。それゆえ、そのような広いエリアを移動中のデータの影響により、計測エラーの検出精度が低下してしまったと考えられる。

修正 Wi-Fi スキャンによる問い合わせによる結果に注目すると、通勤通学データセットにおける精度が悪いことがわかる。通常、Wi-Fi 計測エラーは周囲の AP が少ない場合に発生しやすい。そのため、電車やバスへの乗車により、周囲の AP が少なくなる状況が発生する通勤通学データセットにおいては、この手法は AP のグルーピングが行えないために正確な位置推定が行えない。しかし、周囲の AP が十分に存在する場合には、正確な位置推定を行うことができるため、他のデータセットでは良い結果が得られている。

最後に提案手法に注目すると、他の手法と比べて極めて

表 2 各手法の計測エラーへの分類の詳細結果 (LWL)

Table 2 Classification performance for error detection (LWL)

| | カルマンフィルタ | | | 改良カルマンフィルタ | | | パーティクルフィルタ | | | 他のプロバイダ | | |
|--------|----------|-------|-------|------------|-------|-------|------------|-------|-------|---------|-------|-------|
| | 適合率 | 再現率 | F 値 | 適合率 | 再現率 | F 値 | 適合率 | 再現率 | F 値 | 適合率 | 再現率 | F 値 |
| 計測エラー無 | 0.897 | 0.701 | 0.787 | 0.924 | 0.668 | 0.776 | 0.944 | 0.661 | 0.778 | 0.963 | 0.643 | 0.771 |
| 計測エラー有 | 0.411 | 0.722 | 0.523 | 0.413 | 0.809 | 0.547 | 0.424 | 0.864 | 0.569 | 0.425 | 0.914 | 0.58 |
| 平均 | 0.654 | 0.711 | 0.655 | 0.668 | 0.739 | 0.661 | 0.684 | 0.763 | 0.673 | 0.694 | 0.779 | 0.676 |

| | 修正 Wi-Fi スキャン ($g = 3$) | | | 修正 Wi-Fi スキャン ($g = 5$) | | | 提案手法 | | |
|--------|---------------------------|-------|-------|---------------------------|-------|-------|-------|-------|-------|
| | 適合率 | 再現率 | F 値 | 適合率 | 再現率 | F 値 | 適合率 | 再現率 | F 値 |
| 計測エラー無 | 0.973 | 0.57 | 0.719 | 0.973 | 0.57 | 0.719 | 0.97 | 0.923 | 0.946 |
| 計測エラー有 | 0.388 | 0.944 | 0.55 | 0.388 | 0.944 | 0.55 | 0.772 | 0.901 | 0.832 |
| 平均 | 0.68 | 0.757 | 0.635 | 0.68 | 0.757 | 0.635 | 0.871 | 0.912 | 0.889 |

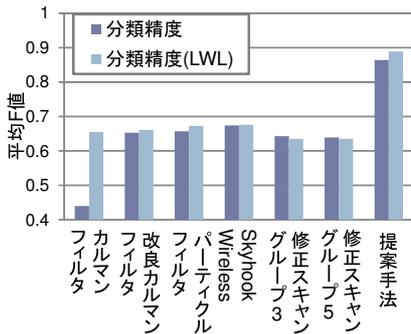


図 6 各手法の分類精度

Fig. 6 Accuracies of naive and proposed methods.

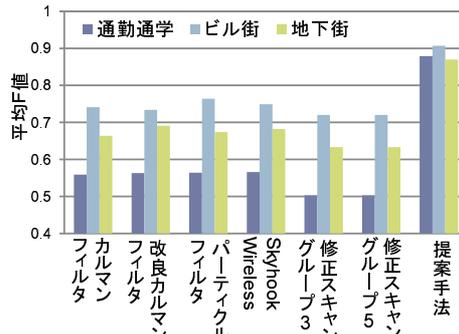


図 7 データセットの各手法の分類精度

Fig. 7 Accuracies of naive and proposed methods for each data set.

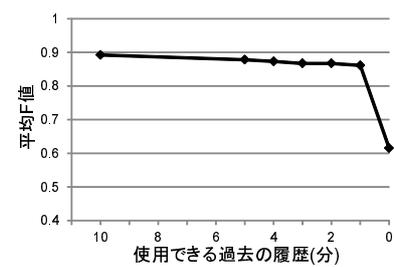


図 8 過去の履歴数による提案手法の分類精度の変化

Fig. 8 Relationship between average F-measure and quantity of history data to be used.

良い精度で計測エラーを検出できていることがわかる。その平均 F 値は 0.889 であり、他手法と比べておおよそ 0.2 も向上している。前述のように、各位置推定器は計測環境に応じてパフォーマンスが変化するが、提案手法は計測環境の特徴量から、最適な位置推定器の意見を計測エラーの検出に反映させることができる。そのため、図 7 のように、3 種類のデータセット全てにおいて、同等の精度を得ることができている。

過去の履歴数による影響: ここまでの実験結果は、すべての過去の履歴が利用可能であるという前提のもとに得られたものである。しかし、実際には Wi-Fi 位置計測時に十分な過去の履歴が存在するとは限らない。例えば、写真等へのメタデータとして位置情報を使用する場合や、ナビゲーションの開始直後では、過去の履歴は存在しないだろう。そのような場合の評価を行うため、使用する履歴を過去 0 分から 10 分まで変化させて、提案手法による計測エラーの検出を行った。その結果を図 8 に示す。本実験では 1 分ごとに Wi-Fi 位置計測を行っているため、履歴の 1 分が 1 組の計測データに対応する。図 8 より、過去の履歴を全く使用しない場合、提案手法の検出精度が極端に低いことがわかる。これは、過去の履歴が存在しなければ使用できない位置推定器や計測環境の特徴があるためである。一方、1 分でも過去の履歴が存在すれば、提案手法は良い検出精度を得ており、これは、前述の過去の履歴が存在しないような場合でも、1 分間待つことで計測エラーの検出が

可能であることを示している。

計測エラーの検出感度: ここでは、提案手法の計測エラーの検出感度について論じる。まず、提案手法がどれほど誤差の小さい計測エラーを検出できるかを検証する。これまでは計測エラーを定義する閾値を 150m として計測エラーの検出を行ってきたが、図 9 に計測エラーを定義する閾値を変化させた時の、提案手法の平均 F 値の変化を示す。図 9 より、閾値を小さくすると検出精度が著しく低下することがわかる。対して、閾値を 150m 以上としても検出精度に大きな変化は見られず、閾値を 150m とすることで問題なく計測エラーの検出が可能であることを確認できる。

次に、提案手法の計測エラー検出における ROC 曲線を図 10 に示す。2 クラス分類器は、対象とするデータがどちらのクラスに分類されるかの確率を算出し、0.5 を超えている方のクラスへと分類を行っている。本実験においては、この分類の閾値を変化させることが、計測エラーの検出感度を変化させることに相当する。ROC 曲線は、この分類の閾値を変化させた際の結果をプロットしたものである。横軸の False positive rate は「計測エラー無」のデータ中で「計測エラー有」と誤分類された割合を、縦軸の True positive rate は「計測エラー有」のデータ中で「計測エラー有」と正確に分類された割合を示している。図 10 から、True positive rate を 0.95 とし、ほぼ全ての計測エラーを検出できるように設定しても、False positive rate が約 0.2 と低い値を保っていることがわかる。

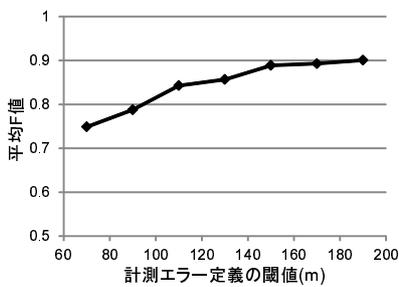


図 9 計測エラーの閾値と分類精度の変化
Fig. 9 Accuracies with changes of threshold of Wi-Fi positioning error.

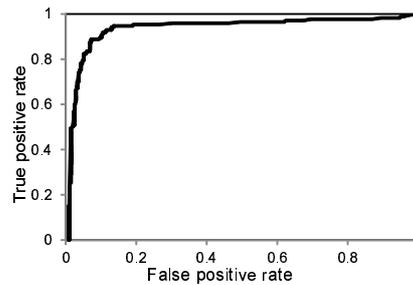


図 10 提案手法の ROC 曲線 (LWL)
Fig. 10 ROC curve for proposed method (LWL).

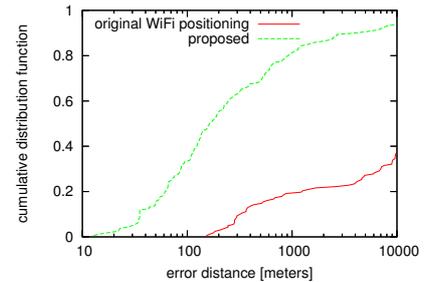


図 11 提案手法による修正前後の誤差の CDF
Fig. 11 CDFs of our method and original Wi-Fi positioning.

以上から、計測環境の変化に対応した計測エラーの検出が可能であり、計測エラーの検出において提案手法が有効であることを確認した。

4.2 Wi-Fi 位置計測における計測位置の修正

評価手法: 評価のために 4.1 節と同様に、leave-one-out 交差検定を行った。計測位置の修正における誤差推定器が使用する回帰分析には、SMO 法を使用し、入力とするデータは誤差が 150m 以上の計測エラーが発生しているデータのみとした。

評価結果・考察: 図 11 は、修正前の Wi-Fi 計測位置と提案手法による修正位置のそれぞれの誤差の累積密度を示している。図 11 における、提案手法による修正位置の累積密度と修正前の Wi-Fi 計測位置の累積密度の違いから、提案手法は修正前の Wi-Fi 計測位置の誤差を多く減少させていることが確認できる。特に、提案手法による修正位置の誤差 150m における累積密度は 0.475 となっている。これは、計測エラーが発生している Wi-Fi 計測位置の内 47.5% を、計測エラーが発生していないと判断される誤差 150m 未満へと修正できたことを意味する。

以上から、提案手法により Wi-Fi 位置計測に用いたデータのみを用いて、Wi-Fi 位置計測の修正が行えたことを確かめ、提案手法の有効性を確認した。

5. 結論

本研究では、Wi-Fi 位置計測において異常 AP により発生する計測エラーを検出し、さらに計測エラー検出時の計測位置を修正する手法を提案した。

Wi-Fi 計測エラーは様々な環境で起こり得るため、単一の位置推定器により正確に現在位置を推定することは困難であった。そこで、Wi-Fi 位置計測を行った周囲の環境を考慮して、様々な特徴を持つ位置推定器を複数組み合わせることで計測エラーの検出を行う手法を実現した。また、同様の位置推定器を用いることで、計測位置を修正する手法を実現した。

評価実験では、実際に様々な状況を想定して収集した Wi-Fi 位置計測データを使用し、提案手法の有効性を検証

した。その結果、計測エラーの検出において極めて高い精度を達成し、計測位置の大幅な修正が可能であることを確認した。

参考文献

- [1] 坂井文泰: GPS 技術入門, 東京電機大学出版局 (2003).
- [2] Vaupel, T., Seitz, J., Kiefer, F., Haimerl, S. and Thielecke, J.: Wi-Fi positioning: System considerations and device calibration, *International Conference on Indoor Positioning and Indoor Navigation*, IEEE, pp. 1-7 (2010).
- [3] 暦本純一, 塩野崎敦, 末吉隆彦, 味八木崇: PlaceEngine: 実世界集合知に基づく WiFi 位置情報基盤, インターネットコンファレンス, Vol. 2006, pp. 95-104 (2006).
- [4] 河口信夫: Locky.jp: 無線 LAN を用いた位置推定とその応用, 信学技報, ITS2007-16 (2007).
- [5] Zhuang, Z., Kim, K. and Singh, J.: Improving energy efficiency of location sensing on smartphones, *The 8th International Conference on Mobile Systems, Applications, and Services*, ACM, pp. 315-330 (2010).
- [6] Lin, K., Kansal, A., Lymberopoulos, D. and Zhao, F.: Energy-accuracy trade-off for continuous mobile device location, *The 8th International Conference on Mobile Systems, Applications, and Services*, ACM, pp. 285-298 (2010).
- [7] Jones, K. and Liu, L.: What where wi: An analysis of millions of wi-fi access points, *IEEE International Conference on Portable Information Devices, 2007. PORTABLE07*, IEEE, pp. 1-4 (2007).
- [8] 藤木慎太郎, 相田仁: ユーザ参加型無線 LAN 位置測定システムにおけるアクセスポイント移設の検出法, 情報処理学会第 72 回全国大会講演論文集. v. 3, p. 302-304 (2010).
- [9] 何韜, 梶克彦, 河口信夫: 位置推定のための無線 LAN 観測データベースの健全性維持手法, 情報処理学会第 73 回全国大会講演論文集. v. 3, p. 281-282 (2011).
- [10] Kang, J., Cohen, I. and Medioni, G.: Continuous tracking within and across camera streams, *CVPR 2003*, Vol. 1, pp. 267-272 (2003).
- [11] Doucet, A., De Freitas, N. and Gordon, N.: *Sequential Monte Carlo methods in practice*, Springer Verlag (2001).
- [12] Google Inc.: Android developers reference: Location-manager, <http://developer.android.com/>.
- [13] Skyhook Wireless: <http://www.skyhookwireless.com/>.
- [14] LaMarca, A., Chawathe, Y., Consolvo, S., Hightower, J., Smith, I., Scott, J., Sohn, T., Howard, J., Hughes, J., Potter, F. et al.: Place lab: Device positioning using radio beacons in the wild, *Pervasive*, pp. 301-306 (2005).