

Wi-Fi 電波強度を利用した機械学習による屋内位置推定における 環境変化への適応手法

濱田航一^{†1} 沼尾雅之^{†1}

概要 : GPS 衛星の電波の届かない屋内でも既存のインフラを用いて低コストで容易に利用できる位置推定手法として、Wi-Fi の電波強度を用いた手法の研究が行われている。中でも、機械学習を用いた Fingerprint 方式は、障害物や反射波などの影響による不規則な電波減衰の発生する環境でも高い精度で推定が行えるという利点がある。しかし、電波強度は周囲の環境や天候、人の出入り、家具の移動など様々な要因によって安定せず、この方式であっても精度の低下が発生してしまう。本研究では、Wi-Fi の電波強度を用いた機械学習による位置推定手法について、誤差を2種類定義し、このうち、環境変化に起因する誤差の軽減を試みた。

キーワード : Wi-Fi 位置推定, 機械学習, Fingerprint 方式

Adaptation Techniques for Environmental Change for Indoor Location Estimation by Machine Learning with Wi-Fi RSSI

KOICHI HAMADA^{†1} MASAYUKI NUMAO^{†1}

Keywords: Wi-Fi position estimation, Machine learning, Fingerprinting

1. はじめに

スマートフォンの普及により、位置情報技術が多く利用されるようになってきている。ナビゲーションや陣取りゲームのような位置情報が必須のアプリケーションのほか、SNS 等でも位置情報を利用した機能が用意されていることがある。商業施設等で現在位置に応じた情報表示や広告配信を行うといったサービスも考案されている。日本では高齢化が進行していることから、高齢者の見守りなどの応用例[1]も注目されている。また、産業分野においては近年 IoT が導入されつつあり、機械や製品の現在地を知ることの重要性が増しつつある。

現在利用されている位置推定技術の代表例として GPS がある。GPS はフィーチャーフォンを含む様々な機器で広く利用されているが、衛星からの電波の届かない屋内では利用することができない。GPS の屋内での代替技術として IMES がある。これは、あらかじめ特定の位置に設置しておいた発信器から発信される位置情報を受け取り、現在位置として利用するものである。電波形式が GPS と同一であるため GPS と併用し屋内外でシームレスに測位を行うことができるが、専用の発信器を導入しなければならず、設置コストが大きい。

そこで、Wi-Fi の電波強度を利用した手法の研究が行われている。この手法は、既存のスマートフォンの機能と Wi-Fi のインフラ設備を利用することができ、システムの構築が安価に行えるという特徴がある。その中でも、機械学習

を用いた Fingerprint 方式は、障害物や反射波などの影響による不規則な電波減衰の発生する環境でも比較的高い精度で推定が行える。

しかし、電波強度は周囲の環境や天候[2]、人の出入り、家具の移動など様々な要因によって安定せず、この方式であっても精度の低下が発生してしまう。本研究では、Wi-Fi の電波強度を用いた機械学習による位置推定システムを構築し、実験によって AP の設置箇所や参照点の周囲の電波到達性などの関係で発生する誤差と家具の移動や人の出入りといった環境変化に起因する誤差の2種類の誤差を定義した。このうち、環境変化に起因する誤差について、2段階の機械学習による手法、AP による合議制を用いた手法、信頼性の低い AP を除外して分類を行う手法の3つの手法を提案し、その軽減を試みた。

2. Wi-Fi の電波を利用した位置推定

2.1 手法の分類

Wi-Fi の電波を利用した代表的な方式としては、大きく電波の到達時間を利用するもの、到達角度を利用するもの、信号強度を利用するものがある[3]。

到達時間を利用した方式は、電波の伝播速度を利用し、到達時間から距離を算出するもので、TOA 方式と TDOA 方式がある。TOA 方式は設置位置が既知の基地局から端末までの到達時間から端末の存在する場所の基地局からの距離を計算し、これを3つの基地局について行うことによって端末の位置を決定するものである。TDOA 方式は複数の基

^{†1} 電気通信大学
The University of Electro-Communications

地局からの端末への伝播の到達時間の差を利用し、端末からそれぞれの基地局への距離の差を二分線として表し、その交点を端末の位置とするものである。これらの方式は精度が高い一方、端末や基地局間で正確な時刻同期をする必要があり、利用は難しい。到達角度を利用した方式は、複数の基地局から端末への電波の到達した角度をもとに、端末の現在位置を推定する方式である。この方式も精度が高いが、電波の到達角度を認識できるアンテナなどが必要になる。

信号強度を利用する方式は、理想的な環境下では電波強度が距離に従って減衰する性質を利用し、基地局からの距離を求め位置を推定する。大きく Proximity 方式、Triangulation 方式、Fingerprint 方式に分けられる。Proximity 方式は、最も信号強度の強い基地局を、そのまま端末の位置とみなす方式である。実装が非常に容易な一方位置の分解能は低く、精度を上げるためには必要な分解能に応じた数の基地局の設置が必要になる。Triangulation 方式は距離による信号強度の減衰モデルを利用し、複数の基地局からの距離によって三角測量のようにして位置を推定する方式である。Proximity 方式より精度は向上するが、電波は障害物や反射波の影響により理想的なモデル通りには減衰しないことが多く、そのような環境下では精度は低下してしまう。Fingerprint 方式（または Scene Analysis 方式）は、推定を行いたい領域内のいくつかの地点であらかじめ RSSI のサンプルを取得しておき、推定時にこのデータと照らし合わせて推定を行う方式である。サンプルを収集するコストはかかるものの、障害物などの影響も考慮した推定ができるため比較的精度が高く、また設置した基地局の位置自体は不明であっても良い。

2.2 関連研究

Wi-Fi の電波強度を利用した位置推定手法の全般的な問題点として、同じ距離や位置であっても、環境変化によって電波の減衰の程度は変化し受信する信号強度は変化してしまうため、精度が低下してしまうということがある。

これについて大野ら[4]は、学習時だけでなく長期的な RSSI 値の収集を行うことにより、電波環境の変化に適応する手法を提案している。この手法では、すべての参照点について長期的に RSSI 値の収集を行うのはコストが高いため、特定の少数の参照点（「長期参照点」）についてのみこれを行い、そのほかの参照点は最も距離の近い長期参照点の RSSI の変化を適用するとしている。この手法ではまた、参照点で電波の観測される頻度が低い AP は安定でない AP として以降の学習・推定の特徴量から除外することにより、モバイル AP のような位置が一定でない AP や遠方にあり安定的に電波を観測できない AP の影響を排除し、精度を向上することができるとしている。

この手法では、長期参照点を設定し常時 RSSI 値の観測

を行わなければならない、長期参照点とするための機器が新たに必要になり、学習データの作成のコストが高くなるという問題がある。また、安定でない AP の除去自体は、移動する AP や電波が弱く推定に利用できないことのある AP を除外するものであり、環境変化に応じて推定に用いる AP を選択するものではない。

何ら[5]は、日本中の無線 LAN 基地局の情報をデータベース化した Locky.jp project について、データベースの健全性を維持し推定精度の低下となる要因を排除する手法を提案している。この手法では、廃棄された基地局、移転した基地局、また新幹線に搭載され位置が一定でない基地局を「異常基地局」と定義し、ユーザーからの情報を利用して基地局情報の修正を行う。

この手法もやはり、電波環境の変化に対応するものではない。また、前提としてユーザーの現在地がわかっている必要がある。

3. 位置推定システム

3.1 位置推定手法

本研究では、Wi-Fi の電波強度 (RSSI) を用いた手法のうち、機械学習による Fingerprint 方式によって位置推定を行う。

機械学習による Fingerprint 方式では、観測された RSSI 値を特徴量とし端末の現在地をクラスとした教師あり学習によって分類モデルを作成し、推定時は観測された RSSI 値を用いて分類モデルによって地点を推定する。横堀ら[6]の研究から、機械学習のアルゴリズムには精度が高く学習時間も短い BayesNet を使用する。

3.2 プローブ要求

Wi-Fi の RSSI を用いた位置推定手法では、AP が端末に対して存在を通知するビーコン信号を利用する方法と、端末が AP を検索するためのプローブ要求を利用する方法が考えられる。ビーコン信号の RSSI を計測する場合、スマートフォン側にパケットキャプチャを行いデータを処理する機構を作成する必要がある、端末所持者に専用アプリケーションのインストールなどを強いることになる。また、端末所持者が能動的にこの機構を起動する必要がある。プローブ要求を利用する場合、端末側は Wi-Fi 接続をオンにしておき、プローブ要求が発信される状態にしておきさえすればよい。また、端末所持者自身が位置推定を行おうとしなくても、位置推定システム管理者側はプローブ要求を収集し、匿名の端末所持者の歩行経路の推定などを行うことができる。

本研究では、スマートフォンから Wi-Fi アクセスポイント (AP) に対して 2.4GHz 帯でプローブ要求を発信させ、AP で観測した RSSI 値を特徴量として用い学習・分類を行う。

3.3 システム構成

システムの概要を図1に示す。

まず推定対象の端末（スマートフォン）がプローブ要求を発信する。APはプローブ要求を受信するとRSSI値をサーバーに送信し、データベースに記録する。このRSSI値に必要に応じて加工を行った上で、学習・分類を行う。本研究では、APは図2の5か所に設置し、学習する地点（参照点と呼ぶ）は25地点とする。

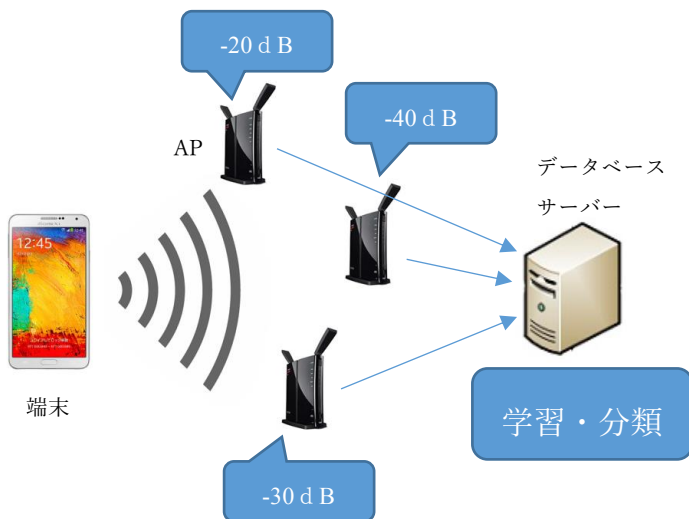


図1 システム概要

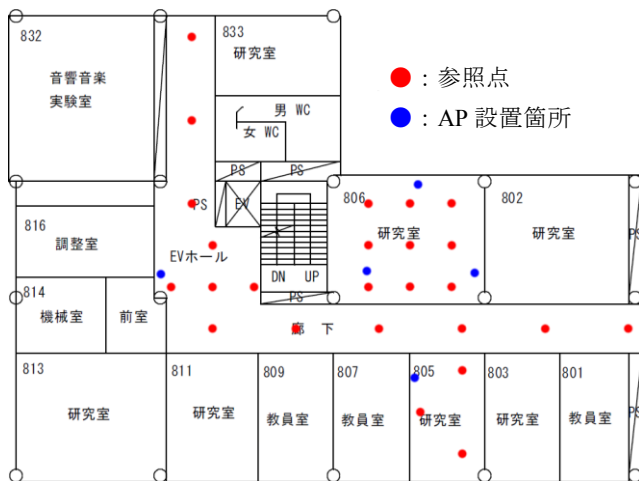


図2 推定領域の見取り図

3.3.1 AP

BuffaloのWZR-HP-AG300H/VにカスタムファームウェアOpenWRTをインストールし、NICをモニターモードに設定する。これにより、Wi-Fiルーターでパケットキャプチャを行えるようになる。推定対象の端末（スマートフォン）から発信されたプローブ要求を受信すると、発信元のMACアドレスとRSSIをサーバーに転送する。

3.3.2 サーバー

APから転送されたMACアドレスとRSSI値から、学習・分類に利用するデータを選び出し、データベースに保存する。

プローブ要求は使用している周波数帯の全てのチャンネルに対して順次発信されるため、APでは待機しているチャンネルの近隣のチャンネルに対する信号も弱い強度で受信してしまう。一般的なスマートフォン端末ではプローブ要求は10秒周期で発信されることを利用し、MACアドレスごとに10秒間で最大のRSSIをデータベースに保存する。これにより、待機しているチャンネルに対する信号のRSSIのみを学習・分類に利用することができる。

3.4 誤差評価

本研究では、スマートフォンの置かれていた参照点と推定結果となった参照点の2次元ユークリッド距離を誤差とする。各実験において、この平均値を誤差評価に用いる。実際の参照点と推定結果の参照点の間には壁などが存在する場合もあるが、ここでは考慮せず一律に直線距離を扱う。

4. 誤差の分析

構築したシステムで実験を行い、誤差を分析した。

まず25の各参照点から5分間ずつプローブ要求を発信し、各30組のRSSI値のデータを得た。別の日程でこれをもう一度行い、結合したデータを利用する。以降の実験でもこのデータを学習データとして用いる。

次に、25の各参照点から1分間ずつプローブ要求を発信し、6組ずつのテストデータを得た。これを別の日程で2回行い、それぞれの25の参照点の平均誤差を比較した。図3に、これを示す。

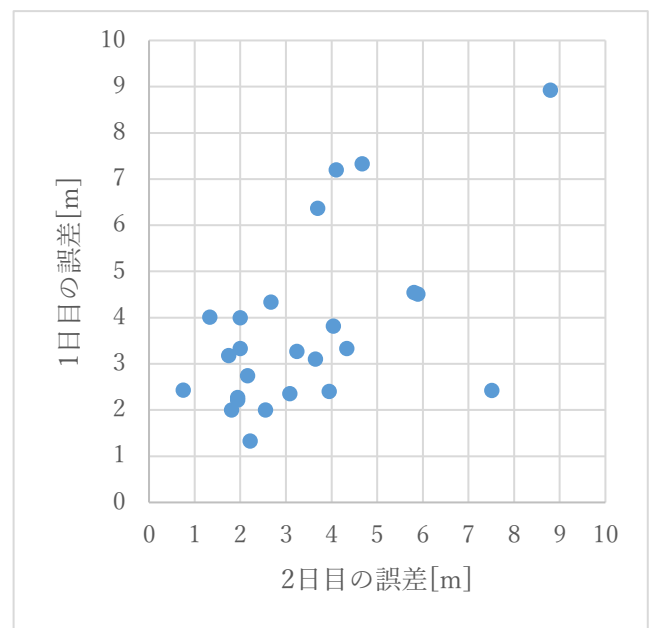


図3 各参照点の誤差

各参照点での両日間での誤差の相関係数は 0.56 で、日による誤差の変動はあるが、日が変わっても参照点ごとに誤差の大小の傾向があることがわかる。この結果から、Wi-Fi の RSSI 値を用いた位置推定における誤差は、「AP の設置箇所や参照点の周囲の電波到達性などの関係で発生する誤差」と「家具の移動や人の出入りといった環境変化に起因する誤差」が存在すると言える。

このうち、AP の設置箇所や参照点の関係によって発生する誤差は、AP の設置数を増やすことによって改善できると考えられるが、その分コストがかかる。また、どこにどのように設置すればよいかは、実際に設置して再度学習・分類を行わなければ判断できない。

以下本研究では、環境変化に起因する誤差について、ソフトウェア上の変更だけで利用できる 3 つの手法によって軽減を試みた。

5. 2 段階の機械学習による手法

機械学習による Fingerprint 方式は、障害物などによる不規則な電波強度の変化も考慮した推定が行えるため、電波の距離減衰を用いて RSSI 値より AP からの距離を求め位置を推定する方式に比べ精度が高い一方、各 AP からの距離自体は全く考慮しないため、観測された RSSI 値によっては極端に大きな誤差が出てしまう可能性が考えられる。

そこで、推定領域を複数のエリアに分割し、まずどのエリアに属するかを推定しさらにそのエリアの中で詳細な位置を推定する方式を用いれば、この極端な誤差を排除できるのではないかと考え、実験を行った。

5.1 提案手法

まず「エリア」を定義するため、各学習時参照点での各 AP の RSSI 値の平均をとり、k-means により学習時参照点のクラスタリングを行う。k は 5 とする。この時分割された 5 つのエリアを図 4 に示す。

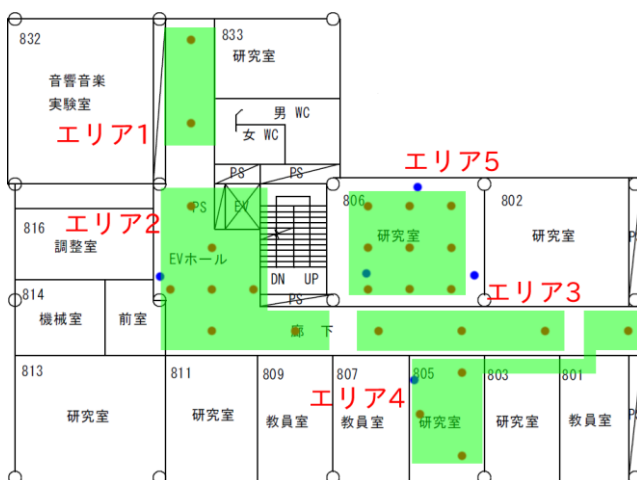


図 4 エリアの分割結果

データ全体から 5 つのエリアに分類する分類器と、5 つ

のエリアの中でそれぞれ参照点を決定する分類器を構築し、テストデータを 2 段階で分類する。

5.2 実験結果と考察

従来手法での全体の平均誤差は 3.20m であったが、提案手法では 3.34m となり、改善は見られなかった。各テストデータの誤差は従来手法と提案手法ではほぼ同じ傾向であったが、この誤差の大きさについて考察する。

まず、1 段階目の属するエリアの推定の正解率は 86% であり、この時点で不正解の 14% は大きな誤差が出てしまうことになる。従来手法で問題になっていた極端な誤差が出ていたデータは、この手法の 1 段階目でも誤って分類されており、結局この問題は提案手法で解決されなかった。

次に、2 段階目で発生した誤差について考える。仮に 1 段階目ですべてのテストデータが正しいエリアに分類されたと仮定し 2 段階目の分類を適用すると、誤差は 2.49m となった。元の 3.20m からは改善されているが、そもそも近隣の参照点との区別が難しく、機械学習による Fingerprint 方式でこれ以上精度を上げるためには、近隣の参照点同士を正確に区別できる必要があることを示唆している。

6. AP による合議制を用いた手法

電波環境は様々な要因で変化し、推定時の RSSI 値を変動させ誤差の原因となる。推定時に電波環境の変化がある AP にその影響があるとき、合議制を用いてこの AP を除外することにより、誤差の軽減を試みる。

6.1 提案手法

まず、5 つの AP の RSSI 値から 1 つ取り除き、4 つの AP の RSSI 値のみを特徴量として学習を行う。これにより 5 パターンの分類モデルを構築しておく。

推定時には、得られた RSSI 値から 1 つを取り除き、対応する分類器を利用して分類結果を出力する。得られた 5 つの推定結果について、平面座標上の XY 座標によって k-means でクラスタリングを行う。k は 2 とする。この結果、5 つの分類結果は以下のいずれかになる。

- 5 つすべての分類結果が同じ
 - 4 つと 1 つのクラスタに分割される
 - 3 つと 2 つのクラスタに分割される
- このクラスタ中心を推定結果とみなし、誤差を検討する。

6.2 実験結果と考察

150 のテストデータのうち、分類結果は以下のようになった。

- 5 つすべての分類結果が同じ：32 ケース
 - 誤差 2.51m
- 4 つと 1 つのクラスタに分割される：61 ケース
 - 4 つの側のクラスタの中心：誤差 3.28m
 - 1 つの側のクラスタの中心：誤差 5.19m
- 3 つと 2 つのクラスタに分割される：57 ケース
 - 3 つの側のクラスタの中心：誤差 3.33m

- 2つの側のクラスタの中心：誤差 4.38m

合議制で、より多数が一致した場合ほど誤差が小さい。5つすべての分類結果が同じだった場合に限っては誤差は2.51mと従来手法の平均の3.20mより小さく信頼度が比較的高いと言えるが、実際にこの手法を利用する場合は合議制によって多数側（5つまたは4つまたは4つの側）を採用することになり、この場合の全体の平均誤差は3.13mとなる。全体では誤差は改善されない結果となった。

7. 信頼性の低い AP を除外して分類を行う手法

前節と同様、信頼性の低い AP を除外する手法について検討する。

AP が人陰に入るなどした場合、同じ地点で10秒差で観測した RSSI 値であっても、その値は大きく異なることがある。そこで、10秒間で RSSI 値が大きく変化した AP を信頼性が低いものとして除外し、残りの AP だけで推定を行う手法を検討する。

7.1 提案手法

学習時には前節と同じ1つの AP の RSSI 値を取り除いたものを利用する。

推定時には、プローブ要求を発信する参照点ごとに無作為に AP を1つ選ぶ。その AP の前に立ったりどいたりし、参照点ごとに2分間ずつプローブ要求を発信する。各参照点で得られた12組の RSSI 値について、偶数組目に以下の操作を行い、これをデータとして分類を行う。

- RSSI 値について欠損値が存在する場合は、これ以上の次元の低下を避けるため除外は行わない
- 欠損値がなければ、各 AP の RSSI 値について直前の奇数組目のものと比較する
 - 比較対象の RSSI 値が欠損値でなければ、5dB 以上の変動があれば除外対象の候補とする
- 除外対象の候補のうち、最も変動値の大きかった AP を除外する
 - 変動値が同率一位の AP が複数存在する場合は、その中から無作為に選んだ1つを除外する
- 変動値が 5dB を超える AP がない場合、比較対象の RSSI 値の中で欠損値があるならば、その中から無作為に選んだ1つの AP を除外する
- いずれにも該当しない場合、除外は行わない

7.2 実験結果と考察

既存手法での平均誤差は 3.89m、提案手法では 4.42m となり、誤差が増加してしまった。

まず誤差が増加した原因として、除外対象の AP の検知率が低かったことが挙げられる。電波環境は常に変化するため、人陰になるようにした AP での RSSI 値の変動が小さく除外対象として選ばれなかった場合や、対象の AP 以外で RSSI 値の変動が起きており、それが優先的に除外対象

として選ばれてしまった場合があった。結果、除外対象の AP を正しく検知できたのは 150 のテストデータ中 56 にとどまった。

そこで、すべてのテストデータで除外対象の AP の検知を正しく行えたと仮定し、AP の除外を行って再度分類を行った。その結果、平均誤差は 3.88m となり、既存手法と変わらなかった。このことから、電波環境の変化が起きる AP がその影響を受けているとしても、それをむやみに除外した推定を行っても、特徴量の次元数が減ることにより精度の向上は見込めないことを示唆している。

8. まとめ

本研究では、まず Wi-Fi の電波強度を利用した位置推定手法で見られる誤差として、AP の設置箇所や参照点の周囲の電波到達性などの関係で発生する誤差と家具の移動や人の出入りといった環境変化に起因する誤差の2種類を定義した。

そのうえで、後者についてソフトウェア上の変更のみで利用可能な3つの手法を提案し、誤差の軽減を試みた。いずれも既存手法に対して優位な結果とはならなかったが、以下の知見を得た。

- 機械学習による Fingerprint 方式で推定精度を上げるためには、近隣の参照点同士の分類が正確にできる必要がある
- AP による合議制を用いることにより、その時点での推定精度の高低を見積もることができる
- 電波環境の変化の影響を受けている AP を除外した推定を行っても、次元数の減少により精度の向上は望めない

これらの知見を活かし、今後推定精度を向上させていく必要がある。また、新規に店舗を構える場合など既存の AP をそのまま利用するのではなく新規に AP を設置する場合を考慮すると、効率よく高い精度で推定が行える AP の設置数・設置箇所の検討も重要である。

参考文献

- [1] 山野太靖, 白松俊, 岩田彰, 永井明彦, Maricio KUGLER. 高齢者徘徊見守りシステムのための Bluetooth と GPS を併用した位置推定手法. 情報処理学会全国大会論文集. 2016, vol. 78, no. 4, p. 995-996.
- [2] “■4 群 (通信工学) -2 編 (アンテナ・伝搬) 11 章 電波伝搬”. http://www.ieice-hbkb.org/files/04/04gun_02hen_11.pdf, (参照 2016-12-18).
- [3] Hui Liu, et al.. Survey of Wireless Indoor Positioning Techniques and Systems. 2007, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 37, NO. 6, p. 1067-2007.
- [4] 大野宇宙, 前川卓也. 電波強度変化を考慮した Wi-Fi 屋内位置推定モデルの低コストな構築手法. 情報処理学会研究報告. 2013, Vol.2013-UBI-40 No.6, p. 1-7.
- [5] 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. 情報処理学会全国大会論文集.

2011, vol. 73, no. 3, p. 281-282.

- [6] 横堀哲也, 沼尾雅之. プローブ要求お利用したスマートフォンユーザー向け屋内位置推定手法. 情報科学技術フォーラム講演論文集. 2014, vol. 13, no. 4, p. 47-50.